# Basenlabile Auxiliare für die Erweiterung der Native Chemische Ligation in Lösung und an der Festphase

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
Im Fach Chemie mit Spezialisierung in Organischer und Bioorganischer Chemie

Eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
M.Sc Ziv Harpaz

Präsident der Humboldt-University zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter/innen:
1. Prof. Dr. Oliver Seitz
2. Prof. Dr. Christian P. R. Hackenberger

Tag der mündlichen Prüfung: 19. November 2015

# Base-Labile Auxiliaries for Extending Native Chemical Ligation in Solution and Solid Phase

DISSERTATION

in partial fulfillment of the requirements for the academic degree of
doctor rerum naturalium
(Dr. rer. nat.)
in Chemistry with specialization in Organic and Bioorganic Chemistry

Submitted to
The Faculty of Mathematics and Natural Sciences
Humboldt University of Berlin

by
M.Sc Ziv Harpaz

President of the Humboldt University of Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Decant of the Faculty of Mathematics and Natural Sciences:
Prof. Dr. Elmar Kulke

Referees: 1. Prof. Dr. Oliver Seitz
2. Prof. Dr. Christian P. R. Hackenberger

Thesis Defense held on: 19th of November 2015

**Selbstständigkeit Dekleration/ Independance Declaration**

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 angegebenen Hilfsmittel angefertigt habe.
Ich habe mich nicht anderwärts um einem Doktorgrad im Promotionsfach Chemie beworben und besitze keinen Doktorgrad im Promotionsfach Chemie.
Die Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 habe ich zur Kenntnis genommen.

I declare that I have produced this doctor's thesis independently using only the tools I have specified, in accordance with section 7 paragraph 3 of the Faculty of Mathematics and Natural Sciences PhD regulations, published in the Official Gazette of Humboldt-Universität zu Berlin (*Amtliches Mitteilungsblatt*) no. 126/2014 on 18/11/2014.
I have not applied for a doctoral degree in the doctoral subject of Chemistry elsewhere and do not hold a doctoral degree in the doctoral subject of Chemistry.
I have taken due note of the Faculty of Mathematics and Natural Sciences PhD Regulations, published in the Official Gazette of Humboldt-Universität zu Berlin (*Amtliches Mitteilungsblatt*) no. 126/2014 on 18/11/2014.

# Table of Contents

# Schemes Index

# Figures Index

# Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Oliver Seitz, for offering a place in your team, for the continuous support during my PhD study, for his motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Secondly, a special thank goes to Simon Loibl who collaborated with me on these projects and his suggestions, ideas and motivation have helped to successfully complete the project. For their helpful and insightful corrections and suggestions to my Thesis I thank Dr. Josephine Schmalisch and Simon Loibl. Dr. Ramon Subiros Funosas and Robert Zitterbart also deserve my sincere appreciations for the helpful advices and discussions. My sincere thanks also goes to my labmates Dr. Alex Roloff, Dr. Melanie Fischbach, Dr. Julia Michaelis and Pablo Lores Lareo for the stimulating discussions, for the nice atmosphere working together, for teaching me German and showing me Berlin. For all the fun we have had in the last five years in and out of the lab. I also thank the rest of the wonderful people in the group for support, good advice and good times together.

Other people who helped make the last 5 years meaningful, interesting and productive are Dr. Rajnish Rao, Saša Bodiroža, Smadar Ovadia-Caro and Mike Ludwig.

Last but not the least, I would like to thank my family: my parents Nili and Chen Harpaz, my sisters Galit Harpaz-Yaniv and Zohar Fenigstein my Aunts Shulamit Persky and Drora Persky and all members of my family for supporting me spiritually throughout the Ph.D and writing this thesis and my life in general.

# 1 Introduction

## 1.1 Proteins and Protein Synthesis

Proteins are the most abundant biological macromolecule and exhibit enormous diversity of biological function ranging from enzymes, hormones, antibodies and transporters to building block and crucial components in organs such as muscle fibers, cytoskeleton or cell membrane. Proteins are the most important products of the genetic information pathway and are omnipresent in all life forms from viruses through single cell bacteria and up to the most complex multi-cellular organisms. They are of major interest in the biomedical, pharmaceutical and material science fields and study of proteins has always been of great interest to the scientific community [1].

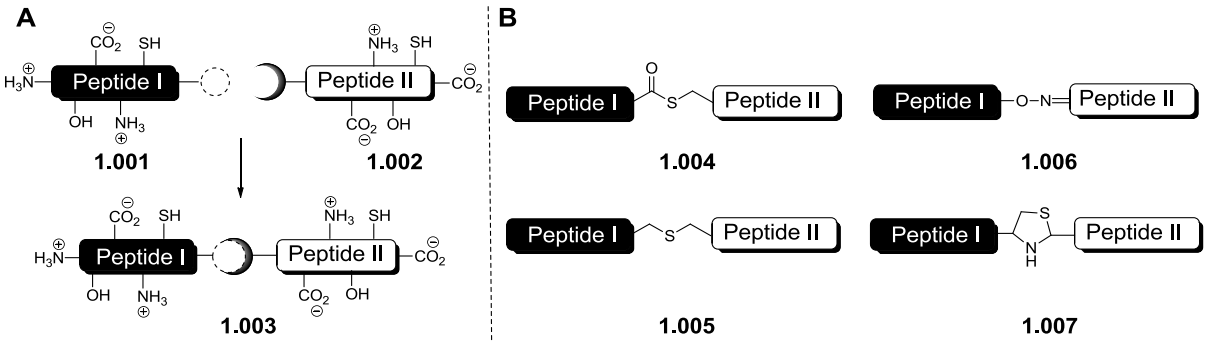
The study of proteins requires acquiring homogenous samples of defined composition, which are then used to determine the structure and function of the protein molecule. Once an understanding of the structure and function of individual protein molecules has been achieved, system-wide studies can be carried out to investigate how proteins interact with one another in the context of the proteome. Nowadays, preparation of proteins is achieved using two main approaches: biological methods to express proteins and total chemical synthesis of proteins.

Current methods in molecular biology typically start with recombinant DNA expression, isolation and purification of a target protein. The structure of the target protein is then systematically varied by modern molecular biology techniques, and the effects of these changes are evaluated by a wide variety of methods aimed to gather new information such as binding affinities, post translation modification, or catalytic mechanisms. Proteins obtained via these methods are restricted to biological constrains, requiring special techniques to allow insertion of unnatural amino acids or modification [2]. Moreover, proteins go through post translational modifications in the cell and proteins of the same sequence can exhibit different modifications, making the protein sample heterogeneous and thus, study of the protein is highly restricted or in severe cases cannot be conducted at all.

Chemical synthesis of proteins is advantageous over the biological methods as it allows total control over the protein sequence. Unnatural amino acids may be incorporated into the sequence selectively [3, 4], acting as probes for structure and function or for therapeutic purposes [5-7]. Moreover, any post translational modification can be introduced in a highly selective manner and in high yields. The previous years have seen development of several chemical ligation methods based on different chemical reactions such as click chemistry[8], oxime ligation[9], traceless Staudinger ligation[10, 11] or ketoacid-hydroxylamine (KAHA) ligation[12]. In the field of chemical synthesis of proteins, chemical ligation methods that rely on acyl-transfer and form the native peptide bond have been established as the paradigm for synthesis.

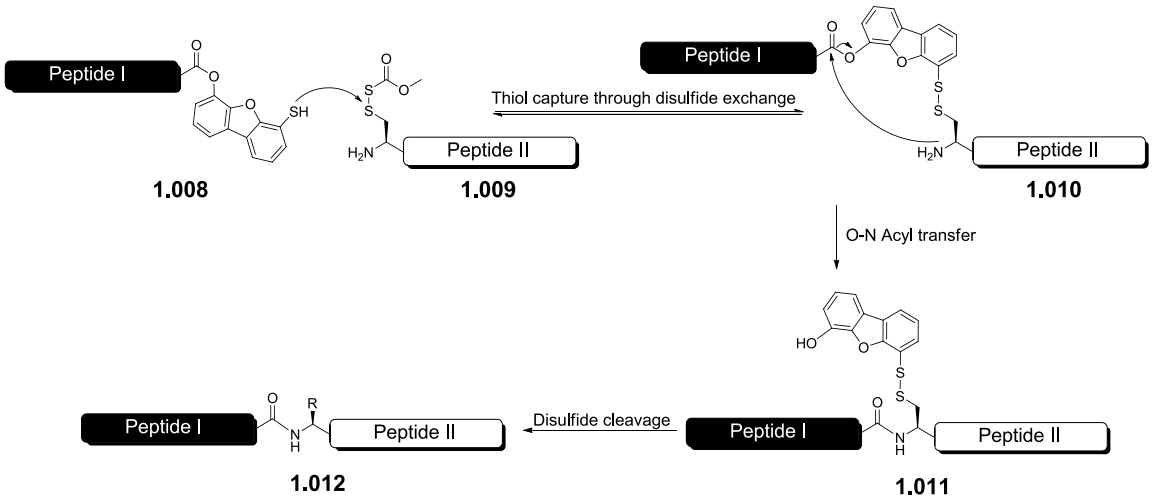
# 1.2 Protein Synthesis through Ligation Methods

Solid phase peptide synthesis (SPPS), introduced in 1963 in the pioneering, ground breaking work of Merrifield [13] is still the method of choice for peptide synthesis. Despite major advancements and optimization in polymer supports, reagents, and automation of the method, SPPS is effective for the synthesis of polypeptide chains of only several dozen residues in length (depending on sequence) [14] due to the accumulation of resin-bound byproducts of incomplete coupling reactions. A number of works have been published presenting impressive synthesis of longer polypeptides using only sequential SPPS but the process is long, tedious, expensive and wasteful (due to multi-coupling steps, use of expensive activators and repetitive washing steps) with low isolated yields [15, 16]. To overcome this limitation, chemical ligation strategies have been developed for the total chemical synthesis of proteins [17]. Chemical ligation enables chemoselective conjugation of two peptide fragments to yield a single defined product. Due to the insolubility of long protected polypeptides, modern chemical ligation methods are performed in aqueous buffers using unprotected peptide fragments to give a larger polypeptide chain in a chemoselective manner (scheme 1.1A). In the early stages, the necessity for chemoselectivity resulted in ligation that formed unnatural bonds such as thioesters, thioether, oxime or thiazolidine [17-20] between the two peptide fragments (scheme 1.1B).




**Scheme 1.1**: **A)** Principle of chemical ligation of unprotected peptide fragments; **B)** non-native chemical ligation bonds. **1.004**: thioester, **1.005**: thioether, **1.006**: oxime, **1.007**: thiazolidine.

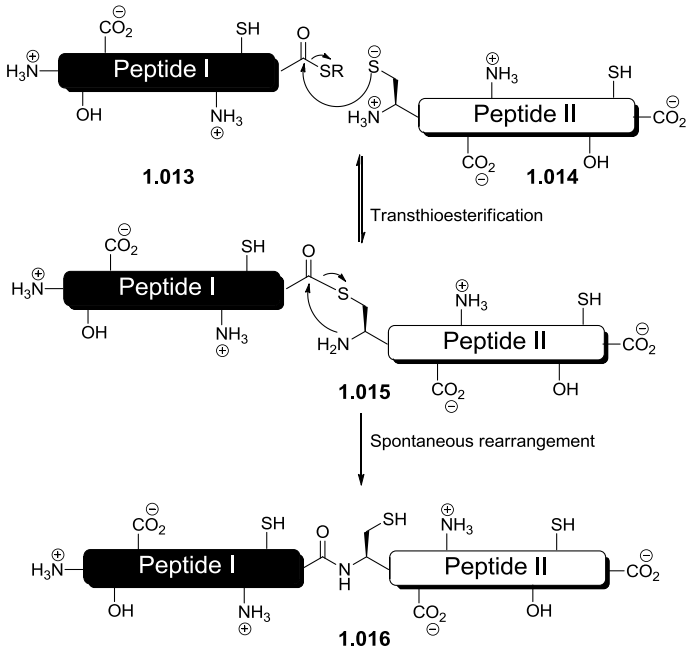
The obvious drawback of the previously described ligation methods was the formation of an unnatural bond at the ligation site forming an unnatural mutation on the target which could limit its study for biological purposes. One of the early successful attempts to ligate two peptides yielding a native peptide bond was achieved by Kemp through prior thiol capture (scheme 1.2) [21]. The method relies on a rigid template based on dibenzofurane equipped with a thiol handle, attached to the peptide C-terminus (**1.008**). In the thiol capture step the disulfide intermediate **1.010** was formed, followed by O→N acyl transfer to yield the ligation product **1.011** which is reduced to the unmodified ligation product **1.012**. The disadvantage is the slow reaction rate due to the medium ring size formed during the acyl transfer (12 member ring), and the poorly reactive C-terminal ester.




**Scheme 1.2**: Peptide synthesis by prior thiol capture.

## 1.2.1 Native Chemical Ligation (NCL)

In 1994, Kent and co-workers developed a new chemical ligation method titled "native chemical ligation" (NCL). Similar to an observation reported by Wieland [22], NCL enabled the formation of the native peptide bond at the ligation junction [23]. NCL involves a terminal $C^{\alpha}$-thioester and a N-terminal cysteinyl peptide. The mechanism comprises two steps: the thioester undergoes a reversible nucleophilic attack by the side chain thiol moiety of the N-terminal cysteine residue followed by spontaneous S→N acyl transfer through a five membered ring transition state, to form a native peptide bond (scheme 1.3). NCL is accomplished in presence of all functional groups present in proteins, in aqueous buffer at neutral pH and ambient temperature. Notably, the chemoselectivity is achieved even in the presence of additional cysteine residues as the formed thioester is reversible and unproductive. Seitz and co-workers have even shown that the presence of internal cysteine residues can accelerate the ligation rate [24, 25]. The first application of NCL was demonstrated by the full chemical synthesis of human interleukine-8 (IL-8), a 72-amino acid chemokine containing four Cys residues. [23].




**Scheme 1.3**: Proposed Mechanism of NCL

## 1.2.2 Extending Native Chemical Ligation

The requirement of a cysteine residue at the ligation junction is an intrinsic limitation of NCL. Cysteine has low frequency in sequences which is estimated to be 1.3% [26]. Some proteins lack cysteine residues entirely, others do not have a cysteine residue at a synthetically beneficial position, and are therefore synthetically inaccessible by NCL. These limitations prompted the development of various methods that seek to extend the high efficiency of NCL.

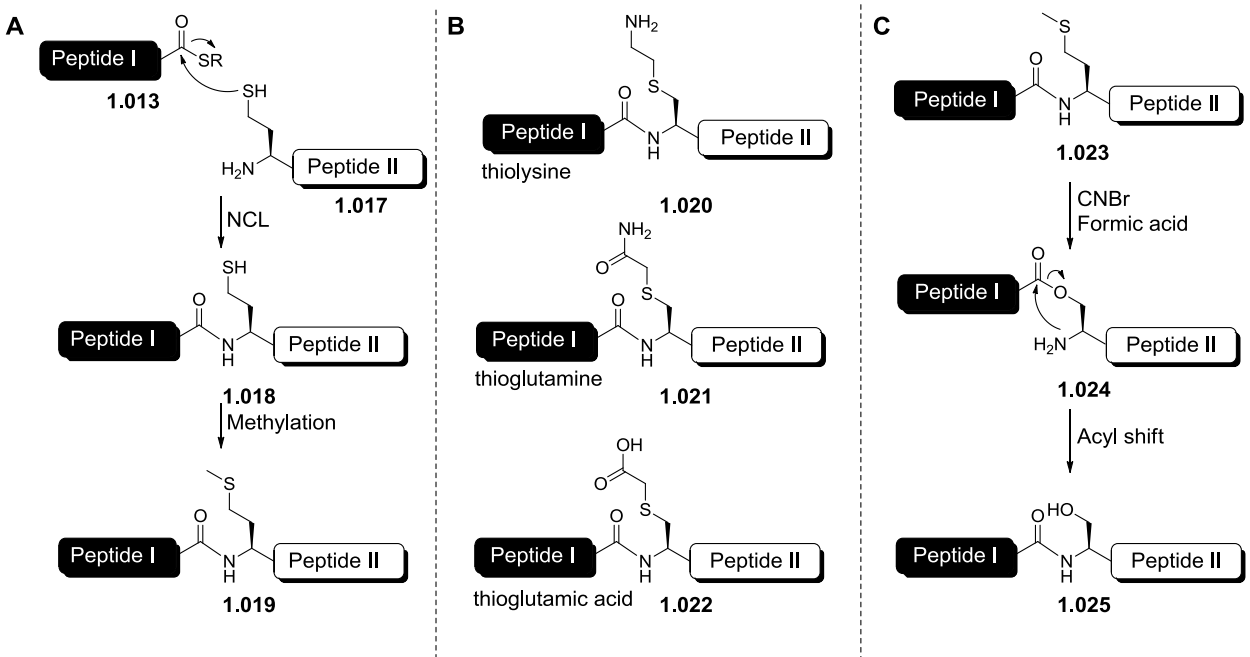
### 1.2.2.1 Conformationally Assisted Ligation

The native chemical ligation of two peptides is achieved by the chemoselective reaction between a cysteine and a $C^{\alpha}$-thioester in denaturing conditions. An interesting approach titled "conformationally assisted ligation" [27] takes advantage of the native tertiary structure of a protein to bring the N-terminal and the $C^{\alpha}$-thioester into close proximity. This allows for the direct attack of the amine on the thioester, eliminating the necessity of the transthioesterification step, and thus, making the cysteine residue at the N-terminal obsolete. Dawson and co-workers demonstrated the effectiveness of the method through the synthesis of Chymotrypsin Inhibitor 2 (CI2), a 64-amino acid protein lacking any cysteine residues making it inaccessible for synthesis via NCL. Early studies have demonstrated that the fragment CI(1-40) of CI2 binds a second fragment CI(41-64) to form a stable protein with a tertiary structure similar to that of the native protein. Initial studies using standard NCL were performed by altering the sequence of the peptide fragments. CI2(1-39)T39D-COSR and CI2(40-64)M40C were synthesized and reacted under ligation condition to give the full CI2 within 8 hr. When

the modified peptide fragments were reacted under folding conditions, the ligation was complete in only 30 min. Ligation under folding conditions proceeded rapidly (30 min) even when the sterically hindered native threonine thioester CI2(1-39)-COSR was used. In comparison, under denaturizing conditions the same ligation required 48 hr to complete. Next, the native CI2(1-39) and CI2(40-64) were ligated. Expectedly, under denaturizing conditions no product formation was detected. When performed under folding conditions product was formed within 2 hr, demonstrating the importance of conformationally induced effective molarity for ligation.

### 1.2.2.2 Thioalkylation and Ligation at Serine Junction

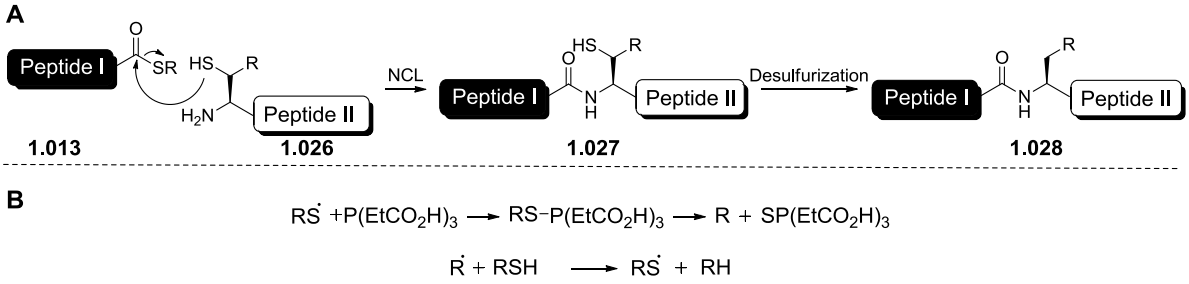
Thioalkylation enables ligation to take place at methionine ligation junctions. In this approach homocysteine (HCY) is introduced into the sequence in place of methionine mimicking the cysteine function in NCL (**1.017**). Following ligation, the HCY is methylated to reform the native methionine **1.019** in the sequence (scheme 1.4A) [28]. Methionine, however, is not common in sequences[26], reducing the applicability for protein synthesis. Moreover, if other cysteine residues are present, protection is needed to avoid alkylation of these residues. By using different alkylation reagents the method can also be used for forming unnatural amino acids at the ligation site such as thiolysine **1.020** [29], thioglutamine **1.021** [29] or thioglutamate**1.022** [30] (scheme 1.4B). The method could also be applied to introduce a serine residue at the ligation site by converting the cysteine upon reaction with cyanobromide to a depsipeptide that subsequently rearranges to form the serine-containing peptide [31] (scheme 1.4C).




**Scheme 1.4:** Extending NCL using thioalkylation. **A)** Ligation at methionine; **B)** Thioderivatives of amino acids; **C)** Insertion of serine at ligation junction following methionine ligation.

### 1.2.2.3 Desulfurization

Desulfurization is a frequently applied method used to convert cysteine to alananine after NCL. The concept has been extended to amino acids beyond alanine. For this purpose, an amino acid bearing a thiol handle at its β– or γ–carbon is required. Following ligation, desulfurization of the thiol handle converts the amino acid into the natural amino acid in the protein sequence (scheme 1.5A).

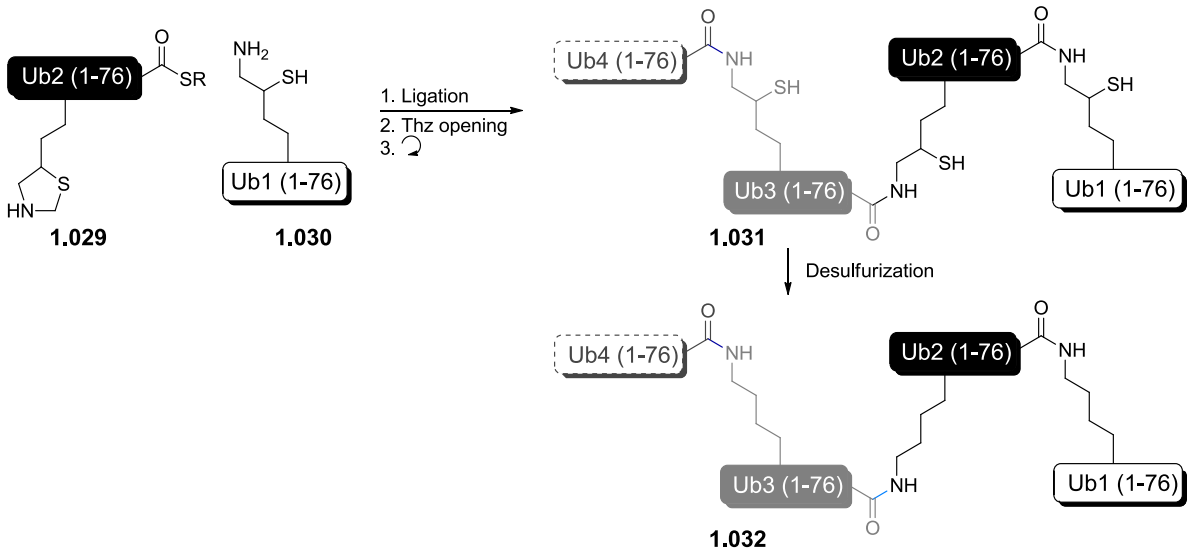



**Scheme 1.5**: **A)** Principle of ligation-desulfurization **B)** Proposed mechanism for free-radical desulfurization.

The method was introduced by Yan and Dawson, who used it to perform ligation at alanine [32]. During synthesis, alanine was replaced by a cysteine and NCL was performed. Following ligation, the cysteine residue was reduced using Raney Nickel to yield the native alanine residue. The application of this approach was demonstrated initially in the synthesis of the cyclic antibiotic microcin J25, the 56-amino acid streptococcal protein G B1 domain and a variant of the 110-amino acid ribonuclease, barnase [32]. The Raney nickel used to induce desulfurization proved incompatible with several functional groups, causing epimerization at threonine and reduction of thioethers and thioesters [33]. A new metal-free desulfurization method was later developed by Danishefsky and co-workers based on a free-radical mechanism compatible with all native functional groups present on peptides (except other cysteine residues that must be protected to avoid desulfurization) [33]. The proposed mechanism involves a desulfurization reaction between sulfhydryl and trialkylphosphine derivatives. A thiyl radical reacts with phosphine, thereby generating an intermediate phosphoranyl radical. Subsequent β scission was envisioned to provide an alkyl radical, and rapid hydrogen abstraction from the parent thiol would furnish the product alkane, thereby serving to propagate the chain (scheme 1.5B). Since its introduction the method gained high popularity and numerous amino acids modified with a thiol handle were developed. Nowadays ligation-desulfurization has been demonstrated with phenylalanine [34, 35], valine [33, 36], threonine [37], lysine [38, 39], leucine [40, 41], proline [42, 43], glutamine [44], glutamic acid [45] aspartic acid [46], tryptophan [4] and arginine [47] junctions.

The method is not without its hindrances though. For one, excluding alanine, γ-proline and penicillamine (ligation at valine) which are commercially available, all of the modified amino acids

must be specifically synthesized. Another disadvantage is the necessity to temporarily protect other cysteine residues to avoid unwanted desulfurization. Thirdly, the ligation product needs to be purified before desulfurization can be performed, as the aryl thiol additives used as catalysts in ligations are excellent radical scavengers and therefore hinder the radical reaction. In recent years several groups have addressed these problems. The group of Brik has used a mercaptophenylacetic acid (MPAA) derivative containing a hydrazine moiety that was removed from the ligation buffer by passing the solution through an aldehyde resin. Desulfurization was sequentially performed without prior HPLC purification [48]. Payne and co-workers have recently demonstrated one-pot ligation-desulfurization using a new thiol additive- trifluoroethanethiol. This highly volatile thiol can be bubbled out of the ligation solution allowing desulfurization to occur [35, 49]. The same group has also published a very interesting study describing one-pot ligation-desulfurization of β-mercaptoaspartic acid without harming other cysteine residues in the sequence, thus avoiding extra protecting group chemistry manipulation and repetitive purification steps [46]. Desulfurization method is credited to some of the most ambitious accomplishments in the field of total chemical synthesis of proteins such as Danishefsky synthesis of the EPO protein [50] or synthesis of the 304 amino acid K64 linked tetraubiquitine employing δ-mercaptolysine to form the isopeptide bond (scheme 1.6) [51].




**Scheme 1.6**: Synthesis of tetraubiquitine isopeptide assisted by δ -mercaptolysine.

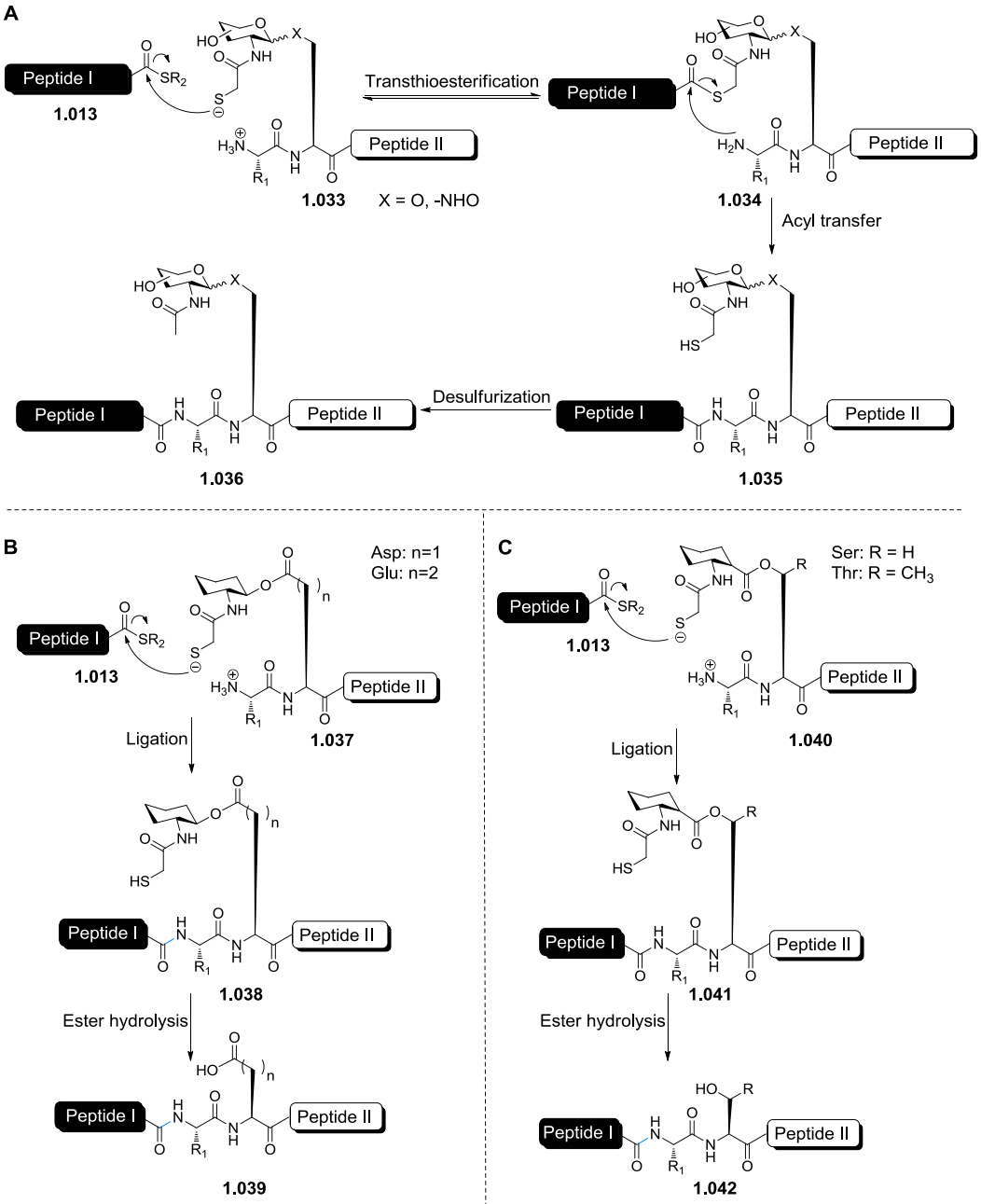
### 1.2.2.4 Sugar-Assisted Ligation (SAL) and Side-Chain Assisted Ligation (SCAL)

Sugar-assisted ligation (SAL) was developed by Brik and Wong to assist synthesis of glycoproteins (scheme 1.7A) [52]. Glycosylated proteins constitute 50% of all proteins in humans. They play central roles in biological processes as protein folding, secretion and cell targeting [53] making glycoproteins highly important targets for biological studies. In this method, a sugar modified with a thiol handle acts as the cysteine residue bringing the glycopeptide **1.033** amine functionality

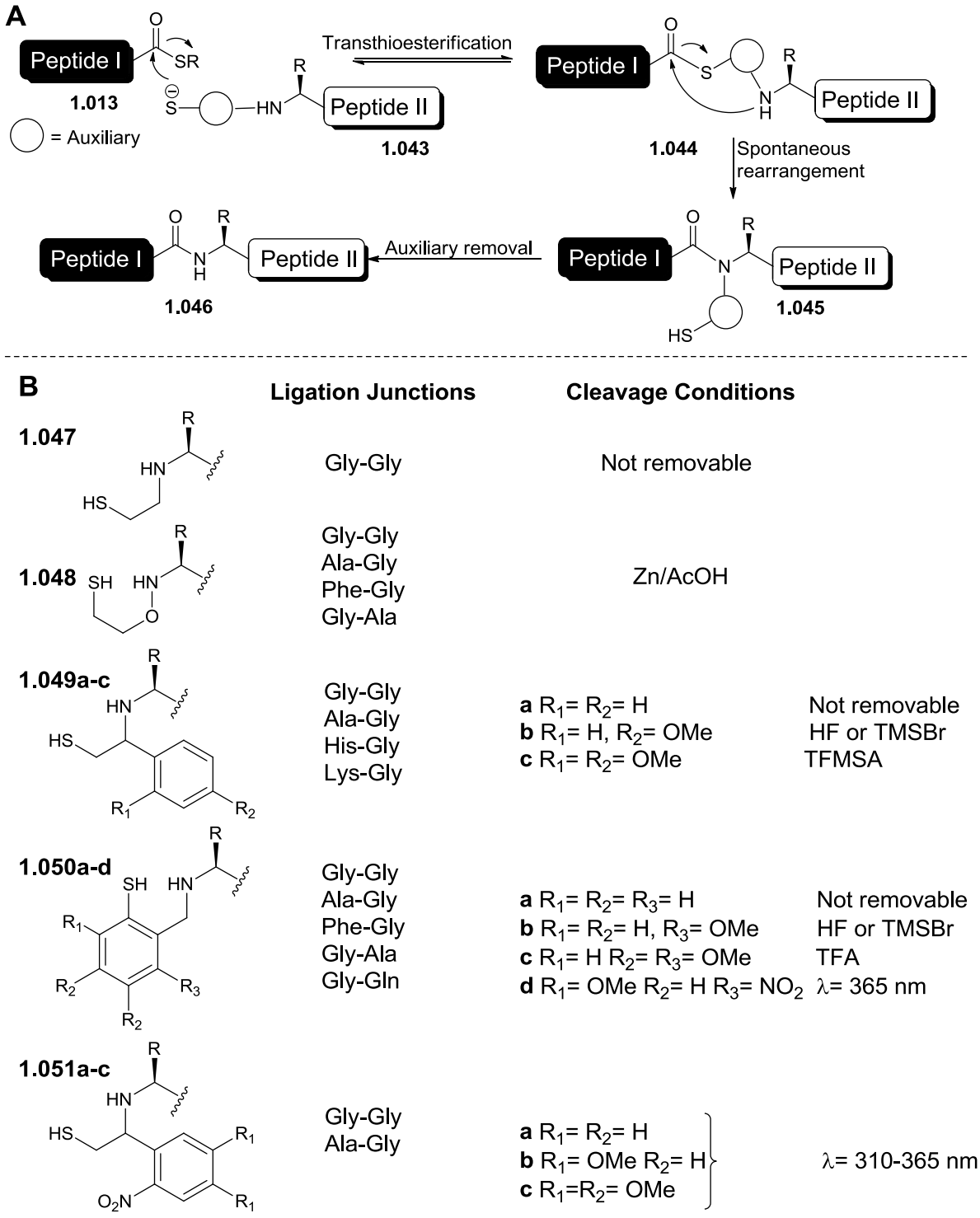
and thioester peptide **1.013** to close proximity and assist the ligation. The thiol handle on ligation product **1.035** is removed using desulfurization methods, yielding the natural glycoprotein **1.036**. SAL was successfully applied to the synthesis of N-, β-O-, and α-O- linked glycopeptides as well as for the total chemical synthesis of diptericin ε glycoprotein [54-57].

Side-chain assisted ligation (SCAL) is a direct evolution of SAL, using a cyclohexane moiety in place of the sugar [58] designed not for the synthesis of proteins. Here, a thiol bearing cyclohexane or a cyclopentane auxiliary, is attached to the side-chain of Asp, Glu (**1.037**), Ser or Thr (**1.040**) residues through an ester bond (scheme 1.7B-C). In a similar fashion to SAL, the auxiliary facilitates the ligation step between two peptides forming the modified ligation products **1.038** or **1.041** respectively. Following ligation, the auxiliary is hydrolyzed to furnish the native proteins **1.039** or **1.042**.

The demonstration for the feasibility of side-chain assisted ligation was shown in the total synthesis of the HIV-1 Tat protein [59]. Tat is an 86-amino acids protein, which plays a crucial role in the life-cycle of the HIV virus. The protein contains seven cysteine residues, but all are located at the N-terminal of the protein thus, the synthesis through NCL cannot be accomplished. The synthesis took advantage of Ser 62 to enable SCAL between the C-terminal peptide and the middle fragment. NCL of the N-terminal peptide furnished the full-length HIV-1 Tat. Difficulties arose when attempts to remove the auxiliary failed. Apparently, once the full protein was achieved, the auxiliary was shielded from the medium and the hydrolysis could not proceed, leaving the auxiliary as an unnatural modification on the protein.




**Scheme 1.7: A)** Proposed mechanism for SAL; **B)** Side-chain assisted ligation at Asp and Glu residues; **C)** Side-chain assisted ligation at Ser and Thr residues.



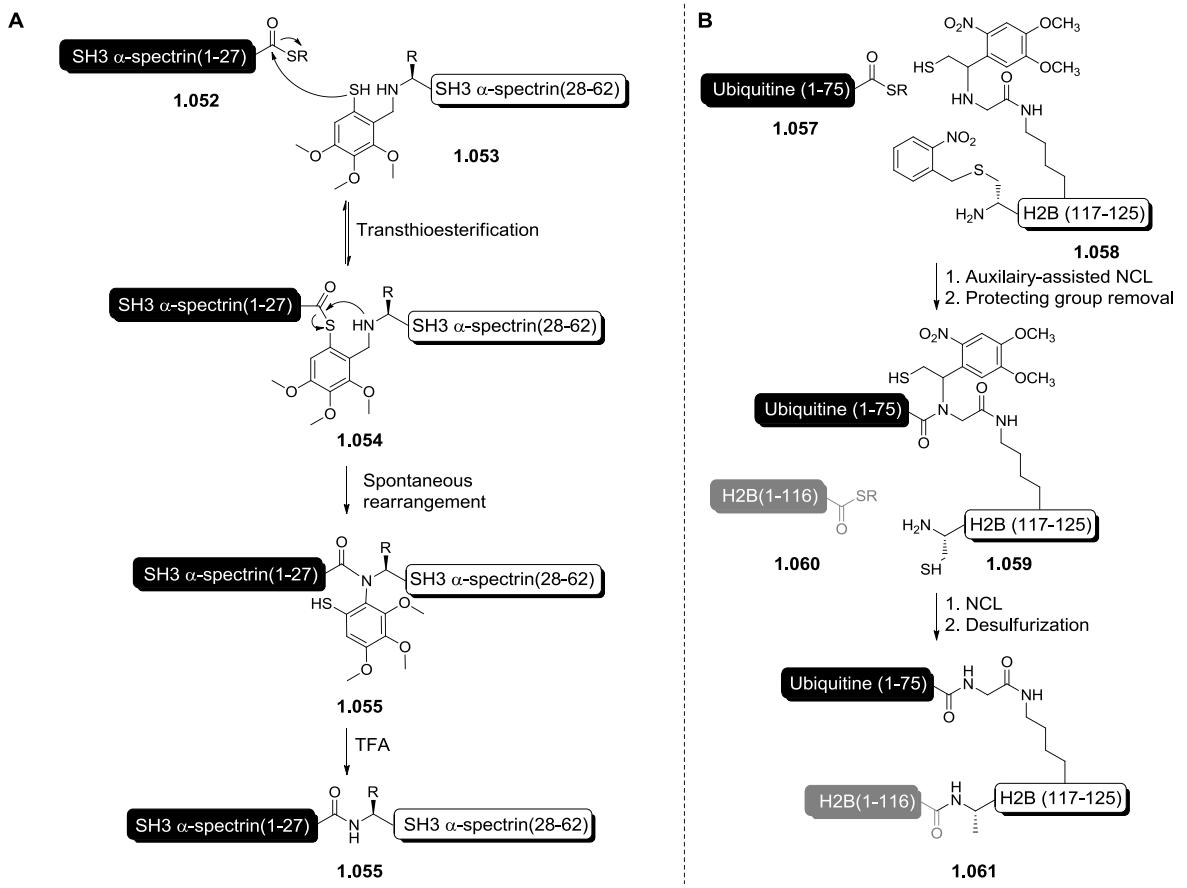

**Scheme 1.8**: **A)** Auxiliary-Assisted NCL; **B)** Ensemble of available auxiliaries and ligation junctions available.

### 1.2.2.5 Removable Auxiliaries

Some of the earliest attempts to expand the scope of NCL used removable auxiliaries. Removable auxiliaries are small molecules bearing a thiol handle that are attached to a terminal amine of a peptide and mimic the cysteine function at the ligation site. Following ligation, the auxiliary is

removed to furnish the natural protein structure (scheme 1.8A). Demands for the auxiliary are an easily accessible synthesis, compatibility with standard solid phase peptide synthesis methods and a chemoselective removal once ligation is complete.

The concept was first reported by Kent and co-workers using $N^{\alpha}$-2-mercaptoethyl auxiliary **1.047** and $N^{\alpha}$-2-mercaptoethyloxy auxiliary **1.048** to promote ligation at a Gly-Gly junction [60]. The mercaptoethyloxy auxiliary **1.048** exhibited significant slower reaction rate in comparison to the mercaptoethyl auxiliary **1.047**, presumably due to less favorable 6-membered ring size of the intermediate. However, the $N^{\alpha}$-2-mercaptoethyl auxiliary **1.047** could not be removed and remained as an unnatural modification. The $N^{\alpha}$-2-mercaptoethyloxy auxiliary **1.048** was removed by treatment with zinc dust under acidic conditions to give the native amide bond at the ligation site. Botti et al. introduced two acid-labile auxiliaries $N^{\alpha}$-1-(4-methoxyphenyl)-2-mercaptoethyl **1.049b** and $N^{\alpha}$-1-(2,4-dimethoxyphenyl)-2-mercaptoethyl **1.049c** [61, 62]. The utility of the auxiliary assisted ligation was demonstrated by Low et al. who used the $N^{\alpha}$-1-(4-methoxyphenyl)-2-mercaptoethyl **1.049b** to synthesize the 106 amino acid protein cytochrome b562 [63]. The auxiliary was later removed by treatment with anhydrous hydrofluoric acid to yield the full-length unmodified product. Offer et al. developed the 2-mercaptobenzyl auxiliary **1.049a**. Although product formation at unhindered ligation junctions was established, the auxiliary was completely stable towards reduction or acidolysis [64]. An improved version of the benzyl auxiliaries was demonstrated first by the group of Aimoto who used the 4,5-dimethoxy-2-mercaptobenzyl auxiliary (Dmb) **1.049c**. This auxiliary was removed post ligation by treatment with trifluoromethanesulfonic acid (TFMSA) [65]. Offer et al. presented an improved auxiliary by introducing the 4,5,6-trimethoxy-2-mercaptobenzyl auxiliary (Tmob) **1.050c** [66]. By increasing the electron density of the aromatic ring in the auxiliary ligation efficiency as well as the acid lability were increased, thus achieving auxiliary removal using the milder triflouroacetic acid (TFA). The auxiliary assisted the synthesis of the 62-amino acid SH3 domain from α-spectrin **1.055** (scheme 1.9A) [67]. Aimoto [68], Dawson [69] and Diederichsen [70] all designed photolabile auxiliaries (**1.051a**, **1.051b**, **1.050d**) by substituting the benzene ring with a nitro group. These auxiliaries however suffer from slow rates and inability to ligate more demanding junctions. Muir used the 2-(4,5-dimethoxy-2-nitrophenyl)ethanethiolbenzyl auxiliary **1.051c** to form an isopeptide bond between a truncated ubiquitine thioester **1.057** and the lysine side chain of a histone **1.058** to examine the biological effect ubiquitination on H2B has on the methylation of Lys79 at H3 (scheme 1.9B). Despite the sterically undemanding Gly-Gly ligation junction, the auxiliary-mediated ligation was extremely slow and was incomplete even after five days [5].



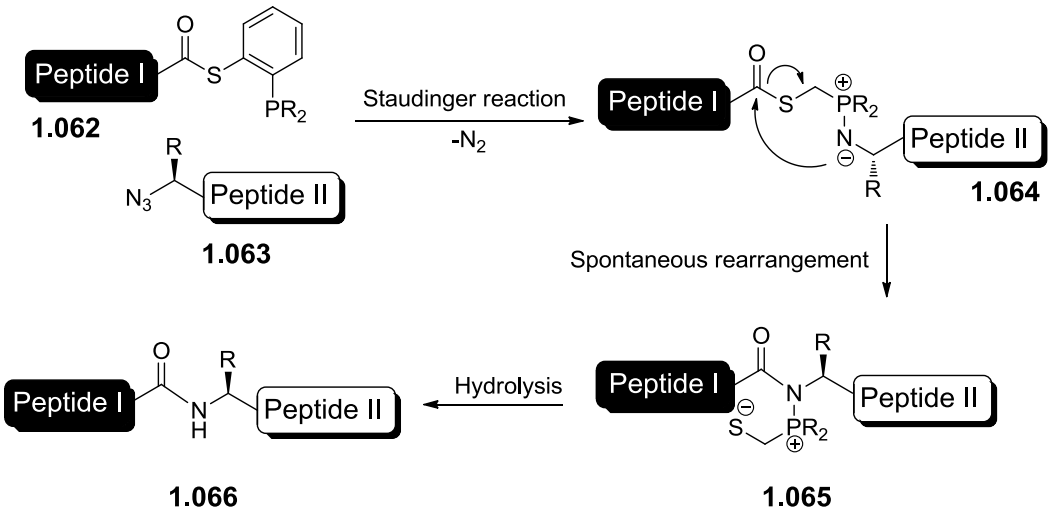

**Scheme 1.9: A)** SH3 α-spectrin synthesis assisted by Tmob auxiliary-mediated NCL; **B)** Isopeptide bond formation between ubiquitine and H2B assisted by the photolabile 2-(4,5-dimethoxy-2-nitrophenyl)ethanethiolbenzyl auxiliary.

The drawbacks of removable auxiliaries are the steric hindrance imposed by the auxiliaries limiting the use of this method to ligation sites containing at least one glycine. A second disadvantage of the acid labile auxiliaries is their incompatibility with acid labile post translational modifications such as glycosylation or sulfonation. Thirdly, under the acidic conditions needed to remove the auxiliaries, the S→N acyl shift reverses back to yield the thioester intermediate thus inhibiting the removal of the auxiliaries from ligation product. This problem can be solved by alkylation of the thiol handle prior to acid treatment to prevent the N→S acyl shift. However, it requires other cysteine residues in the sequence to be protected in order to avoid unwanted protein modification.

## 1.2.3 Traceless Staudinger Ligation

The reaction was introduced by Bertozzi[10] and Raines[11] and is based on the known Staudinger reaction[71], in which an azide reacts with a phosphine to form aminophosphorane. In the traceless Staudinger ligation the formation of the iminophosphorane **1.064** brings the nitrogen atom in close proximity to the thioester leading to an acyl shift and forming the amidophosphounium salt **1.065**. The complex is consequently hydrolyzed to yield ligation product **1.066** (Scheme 1.10). Since its introduction several aspects of the Staudinger traceless ligation have been improved: water soluble phosphinothiols have made it possible to use thioesters and form the phosphothioesters in-situ[11, 72].

Raines also reported ligation at non glycyl azides. It was used for the synthesis of RNAse through the condensation of three segments RNAse A(1-109) was expressed as a thioester, the dipeptide RNAse(110-11) was synthesized as the phosphinothioester and RNAse(112-124) was synthesized as the N-terminal azide. The Staudinger ligation was performed directly on solid support after which the peptide was cleaved and purified before performing the final NCL to give the full-length RNAse[73]. It was also used for synthesis of cyclic proteins by protecting the phosphinothiol with borane which allowed the azide group to be added on the same molecule[74]. After protecting groups removal the ligation was performed. Lee et al. developed a resin bound phosphorylthiol for the solid phase Staudinger ligation. The resin can be recycled for repetitive use.



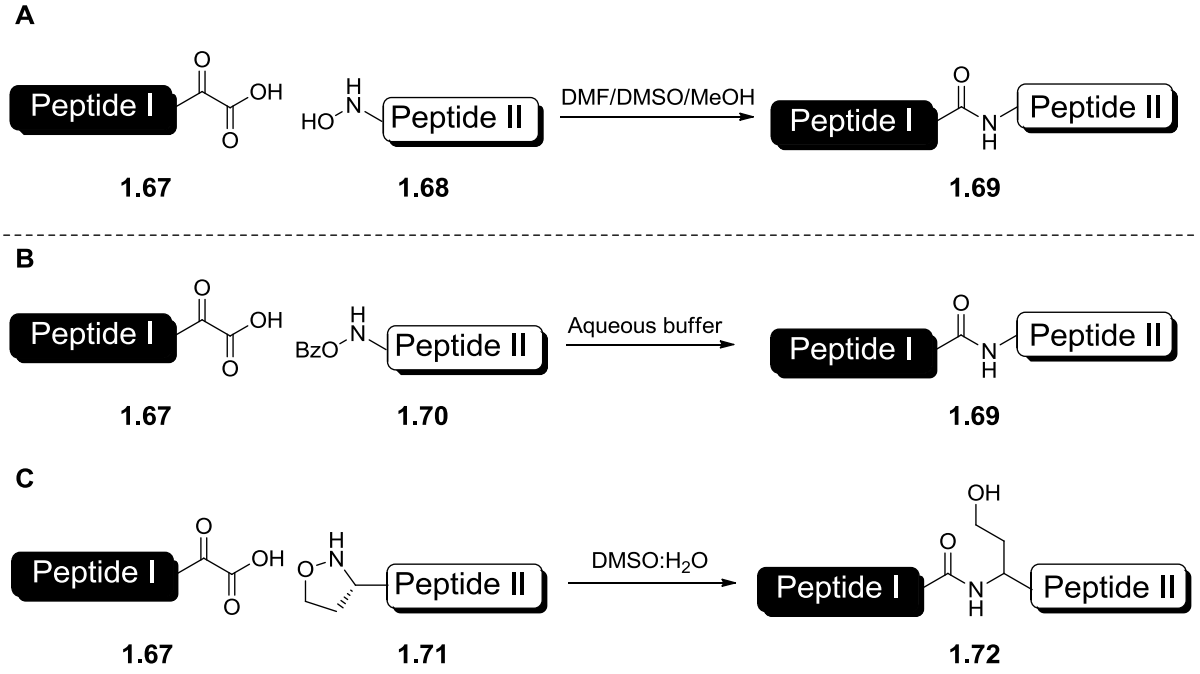

**Scheme 1.10**: General mechanism of traceless Staudinger ligation

## 1.2.4 Ketoacid Hydroxylamine Ligation

Ketoacid-hydroxylamine (KAHA) ligation was developed by Bode et al as an alternative method for reagent free peptide bond formation. The method is based on the selective reaction between α-ketoacid and hydroxylamine to yield an amide bond in presence of other functional groups often found in peptides such as amines, alcohols and carboxylic acid [12]. Two types of KAHA ligations with two distinct pathways were defined: type I ligation (scheme 1.11A) occurs between peptide α-ketoacids and O-unsubstituted hydroxylamines in polar aprotic solvents (such as DMF or DMSO) and peptide concentrations comparable to NCL. Type II ligation (scheme 1.11B) requires O-substituted hydroxylamines (mostly as O-benzoyl hydroxylamines) and works best at aqueous medium and higher peptide concentrations. The benzyl substituted hydroxylamine undergoes elimination making their use for ligation problematic. A solution to the labile benzoyl protected hydroxylamine came in the form of 5-oxaproline. This building block can be introduced directly during SPPS and reacts selectively with ketoacids to give the peptide bond at the ligation site and an unnatural

homoserine residue (scheme 1.11C) [75]. Using oxazetidine building block as alternative to 5-oxaproline also proved successful and allowed the introduction of a serine residue at the ligation site[76]. Major advancements in KAHA ligation were the development of SPPS compatible synthesis of the peptide α-ketoacids and hydroxylamines [75, 77], the identification of 5-oxaproline as a stable hydroxylamine and the ability to mask the ketoacids and hydroxylamines to enable multiple segments ligation. The ligation methods have been successfully applied for synthesis of several protein targets such as the Pup[75] UFM1 [78] and heme protein Nitrophorin 4 [79].

Although impressive demonstrations of the method capabilities, KAHA ligation remains a non-common practice mainly due to the requirement of DMSO as co-solvent, the somewhat elevated temperatures required for ligation of longer peptide fragments, introduction of the unnatural homoserine residue as well as the need to purify the ligation products before unmasking of the ketoacid and hydroxylamine.



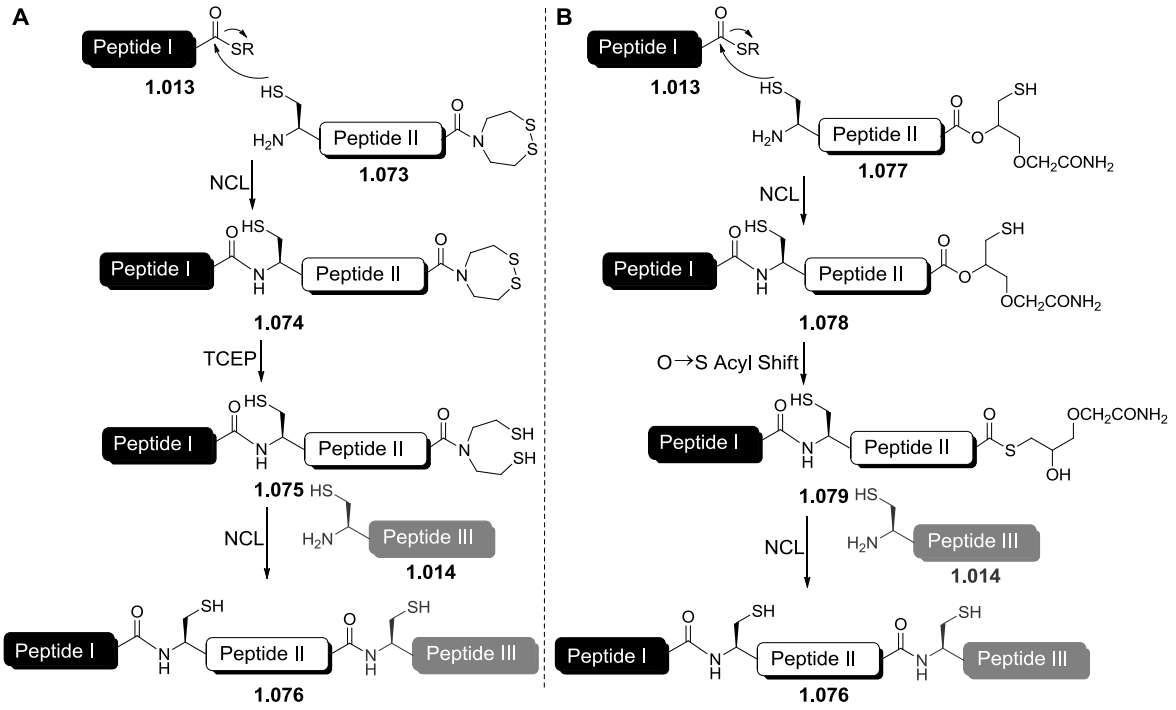

**Scheme 1.11**: Ketoacid-hydroxylamine ligation. **A)** type I KAHA ligation; **B)** type II KAHA ligation; **C)** KAHA ligation employing 5-oxaproline as hydroxylamine derivative.

## 1.2.5 Solid Phase Chemical Ligation (SPCL)

Most proteins and peptides synthesized nowadays using chemical ligation methods are synthesized in solution. As the ligation chemistry is done in aqueous buffers under ambient temperatures, working in solution is comfortable as it allows easy and immediate analysis of the ligation by chromatography techniques combined with mass analysis and direct purification by preparative HPLC with no work-up required prior to it. Despite the advantages of NCL in solution, it is still a limited tool for synthesis of larger targets of several dozen kDa in size (largest protein

synthesized by pure chemical methods was the 304 amino acid tetraubiquitine). In early years after NCL introduction, such an endeavor required purification of the product after each step, not only drastically decreasing the yield but also being time consuming. In the recent decade, major advances in multi-segment sequential and convergent ligations have been made. Kent and co-workers have demonstrated a one-pot ligation of 3 segments of crambin using a thiazolidine (Thz) protection on the cysteine of a middle fragments [80]. The Thz protecting group has enjoyed broad use in multi-segment ligation due to its mild removal conditions and compatibility to ligation conditions [40, 81-83]. A different approach relies on masked or latent thioesters derivatives that are activated post ligation to enable consequent ligations. Examples include the bis(2-sulfenyl)amino (SEA) linker **1.073** developed by Melnyk [84-86] (scheme 1.12A) or Kent's kinetically controlled ligation (KCL) [87]. Other examples utilizing controlled O→S (scheme 1.12B) or N→S acyl shifts have also been presented [41, 88-91]. These attempts, while extending the scope of chemical synthesis of protein, have yet to offer a general, universal solution for sequential ligations.

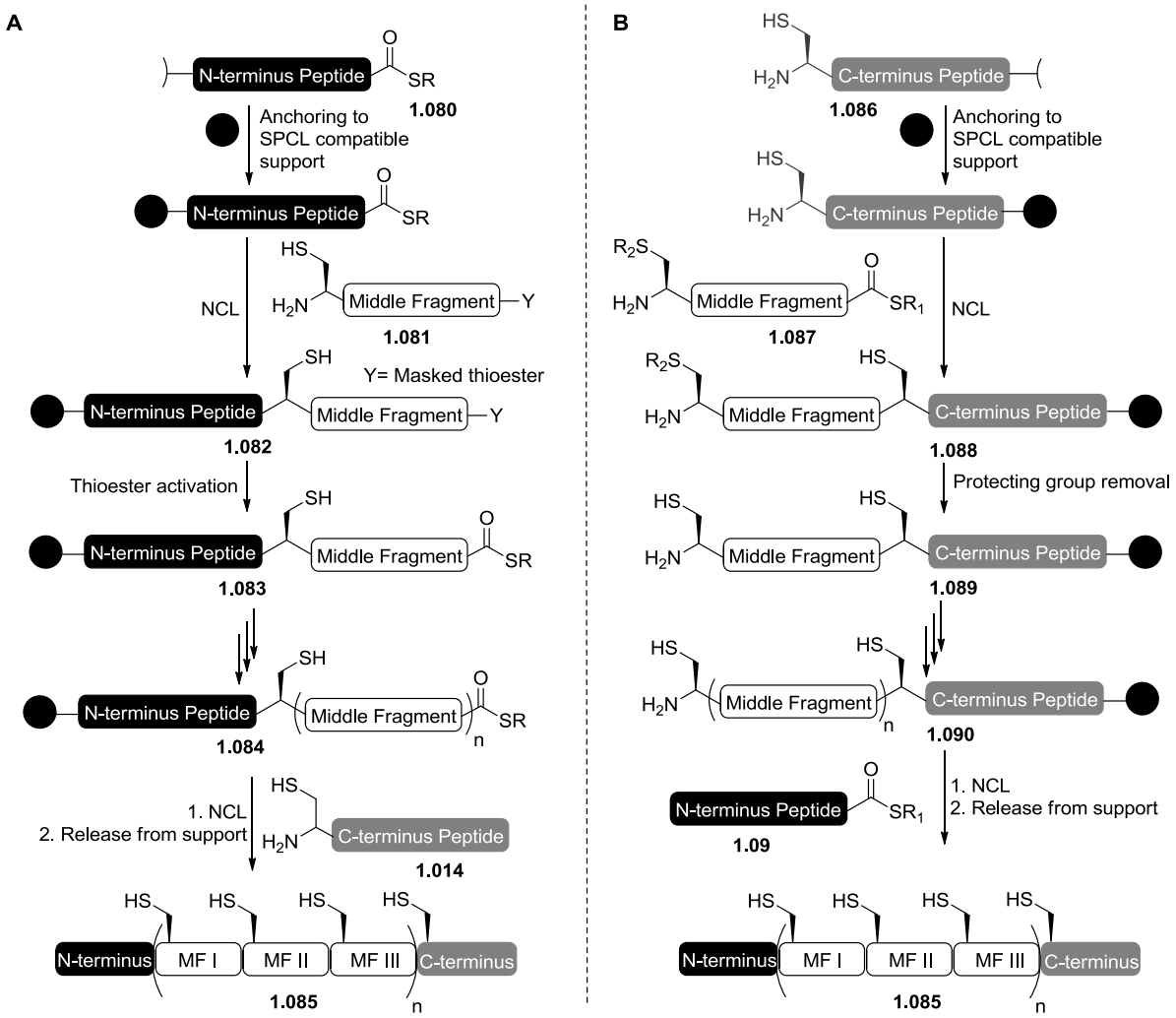
Having a general tool that enables several ligations to be performed sequentially would increase the number of targets achievable through chemical synthesis. For this reason, trials were made from very early on to adapt NCL to the solid support medium. The motivation behind these trials is clear: utilize the advances SPPS has given to peptide synthesis for the total chemical synthesis of proteins. By successfully applying NCL to the solid support one could use large excess of reagent to force ligations to completion and easily remove them by simple washing steps, perform consecutive reaction, and possibly automatize the process. The principles of SPCL are shown in scheme 1.11. A target protein can be assembled by sequential ligation steps on solid support in either N→C or C→N direction. First, peptides need to be orthogonally and selectively anchored onto and released from the solid support. The direction of the synthesis determines the property of the middle fragments. In N→C directed SPCL (scheme 1.13 A) thioesters precursors are utilized (**1.081**).



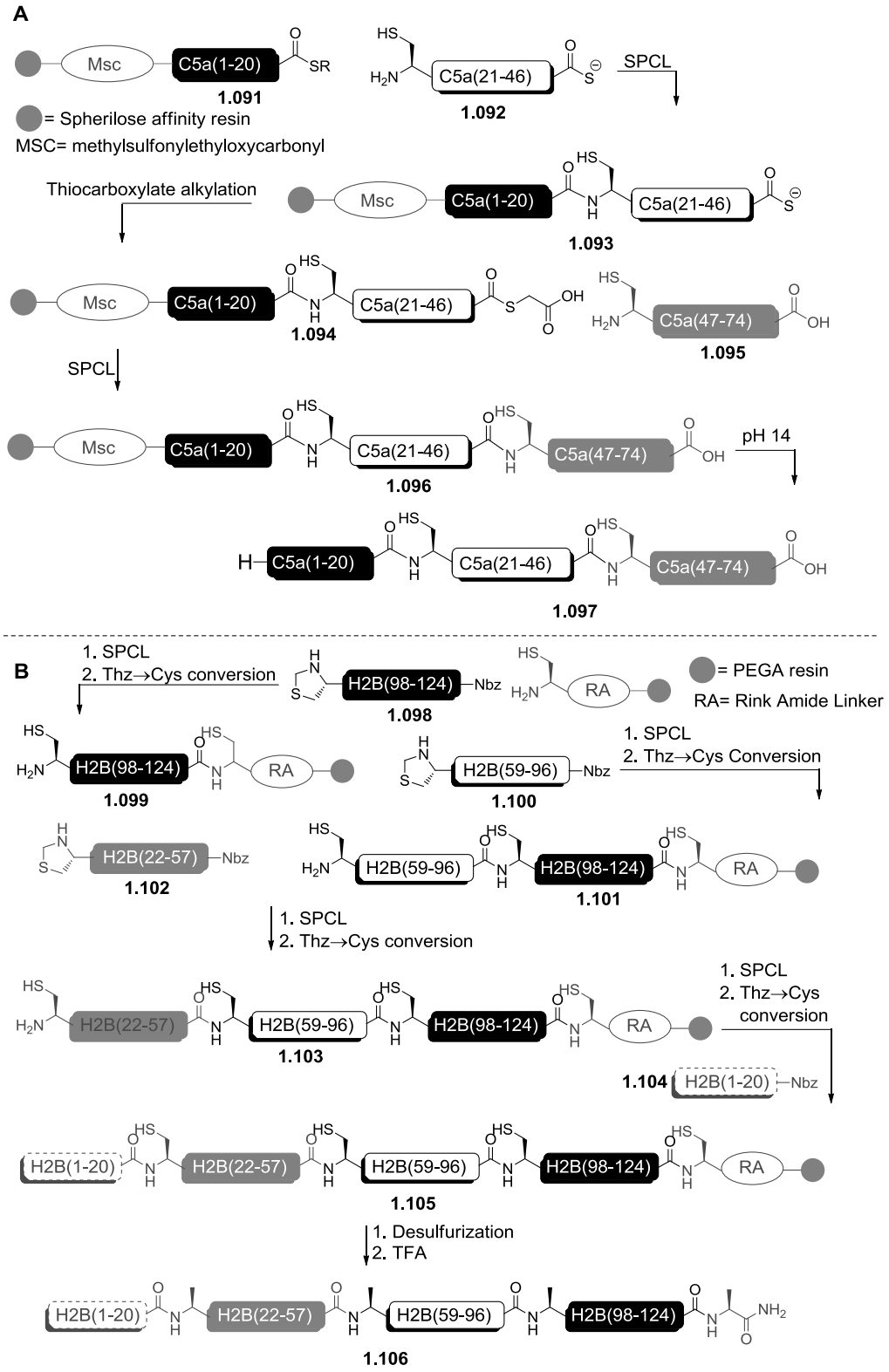

**Scheme 1.12**: N→C directional three-segment NCL utilizing crypto thioesters. **A)** SEA$^{on/off}$ system, **B)** Ligation of peptide O-Esters.

These are activated post-ligation to enable chain elongation (scheme 1.13A **1.083**). If the C→N direction is chosen, the middle fragments N-termini must be orthogonally protected to allow controlled ligation (scheme 1.13B).

Indeed, several pivotal works have shown the potential of SPCL as a method to synthesize proteins with good yields and purity. The group of Kent has released the first strong evidence for the methods applicability, synthesizing three proteins of varying length: the 74 amino acid long human anaphylatoxin C5a (scheme 1.14A), the 115 amino acid long human MIF protein and the 118 residues long GV-PLA$_2$ protein. Syntheses were demonstrated in both N→C as well as C→N directions [92]. Melnyk used the SEA linker for SPCL application to synthesize the 136 amino acid long polypeptide derived from human hepatocyte growth factor by ligation of 5 segments in a linear fashion [93]. Recently the group of Brik has successfully utilized SPCL-desulfurization to synthesize the 124 amino acid long H2B protein by condensation of four peptides (scheme 1.14B) [94]. The cysteine residues were masked as thiazolidines for selective coupling and following the complete assembly of the protein, the unnatural cysteine residues were desulfurized directly on support to yield the native alanine residues.




**Scheme 1.13**: General scheme for sequential SPCL. **A)** N→C direction by using masked thioesters; **B)** C→N direction relying on orthogonal protection.




**Scheme 1.14**: **A)** Synthesis of the C5a protein in N→C directional sequential SPCL using thiocarboxylate as masked thioester. **B)** Synthesis of H2B protein in a C→N directional sequential SPCL utilizing Thz as masked Cys and on resin desulfurization. Msc: methylsulfonylethoxycarbonyl, RA: Rink Amide.

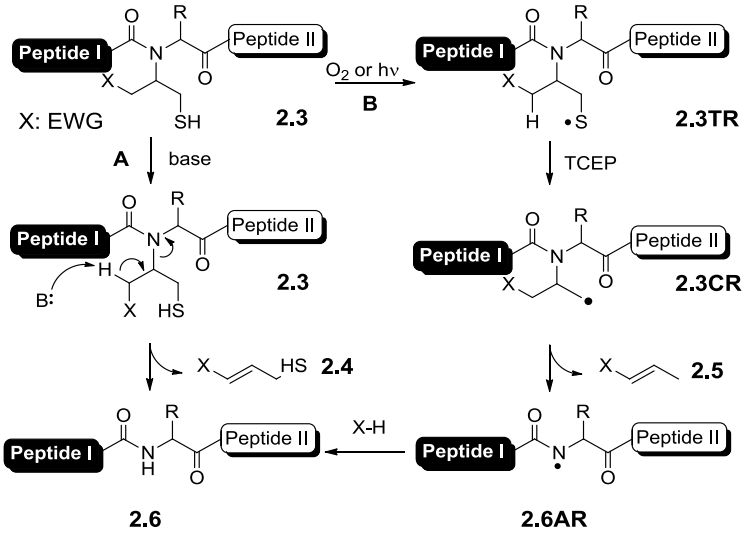
# 2 Research Goals

Several methods for extending NCL were developed over the recent years and have expanded the scope of NCL greatly. As a result, many new ligation junctions were introduced, giving great flexibility for chemists in the field of total chemical synthesis of proteins. Despite the numerous ligation junctions available, the goal of enabling NCL at any residue, which would give complete synthetic control and access to a wide variety of proteins, is not yet achieved. Moreover, some of the methods that were developed do not exhibit the same effectiveness as NCL in terms of ligation time and yield, other methods, such as conformationally assisted ligation are extremely specific and are useful only in rare cases. The desulfurization technique, although very promising and despite the great achievements accomplished through it, enables ligation only at a specific junction and often requires laborious synthesis to yield the thiol-handle modified amino acids..

The goal of this research was to develop an effective method for peptide ligation that could be broadly used for total chemical synthesis of proteins in solution and on solid-support. Auxiliary assisted ligation methods have great potential in realizing this vision. An auxiliary can be attached to the N-terminus of any peptide, promote ligation at wide range of junctions, and removed to yield the native protein. However, the biggest hindrance encountered with auxiliary-assisted NCL was the inability to ligate sterically demanding junctions efficiently. The harsh conditions needed to promote auxiliary removal from the ligation product is another factor preventing wider application of the method as some of the auxiliaries require strong acids such as hydrofluoric acid, TFMSA or TFA which are dangerous to handle or are incompatible with frequently occurring post translational modifications. To overcome these limitations, an innovative auxiliary that could be effectively and easily attached to peptides during SPPS and allow ligation at a broad range of ligation junctions was sought after. The auxiliary was designed to be removed from the amide bond following ligation by exposing the ligation product to mild basic conditions in aqueous medium and should be compatible with unprotected proteins and acid-sensitive post translational modifications.

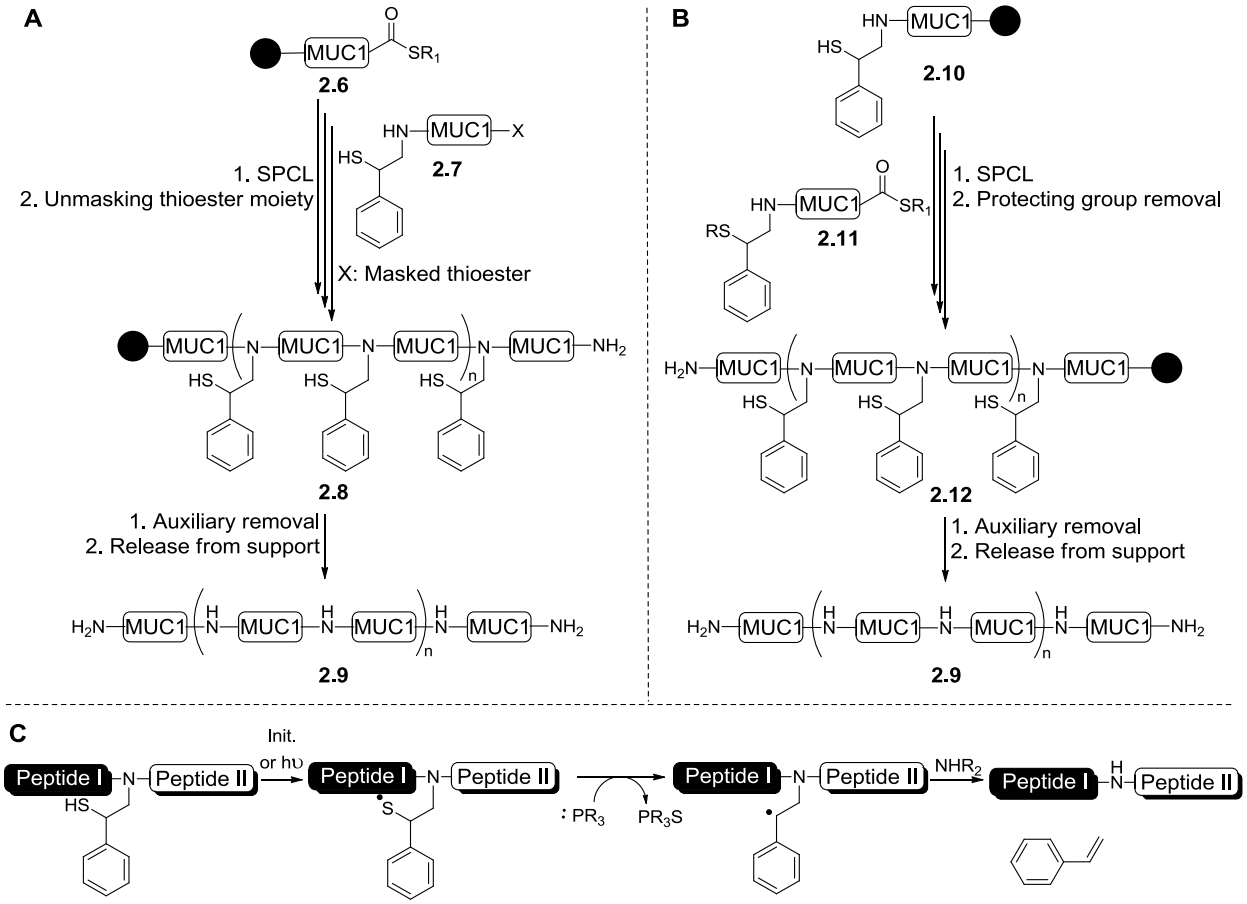
In the first part of the work development of such class of auxiliaries is presented and discussed. The design and hypothesized mechanism of these auxiliaries is described in scheme 2.1. The auxiliaries are based on the 4-mercaptobutanoate attached to the terminal amine via C3. The thiol handle is positioned in such a manner as to form the favorable 5-membered ring intermediate, mimicking the cysteine role in NCL. The linear butyl scaffold should provide flexibility to the auxiliary thus enabling it to promote ligation at hindered junctions. The carboxyl function presents several electron withdrawing substituents which increase the acidity of the α-proton. Once a peptide bond is formed, cleavage of the auxiliary via elimination was envisioned to take place by exposure to mild basic conditions furnishing the native peptide bond and a mercaptobutenoate derivative (scheme 2.1 path A).

During the course of the ongoing PhD work of Simon Loibl, Humboldt University Berlin, observations were made that hinted towards auxiliary cleavage by a radical mechanism. In agreement with previously suggested radical desulfurization, exposure to oxygen or light could lead to formation a thiyl radical **2.3TR**. In presence of TCEP, desulfurization may take place to yield the carbon radical **2.3CR**. Fragmentation may lead to the butenoate derivative **2.5** and peptide amide radical **2.6AR** which would abstract a hydrogen from a hydrogen donor (an amine base) to yield the native peptide (scheme 2.1 path B).




**Scheme 2.1:** General scheme depicting the auxiliarystructure and hypothesized removal mechanisms. A) ionic mechanism; B) radical mechanism.

In the second part of the work the concept of auxiliary assisted NCL was expanded further for the first reported application of auxiliary-mediated SPCL using a 2-mercapto-2-phenylethyl auxiliary peptide **2.7** (scheme 2.2 A and B)[95]. The auxiliary was developed by Simon Loibl (Humboldt University Berlin) on in the work towards dissertation. Similarly to the other auxiliaries, ligation undergoes a favorable 5-membered ring transition state as to benefit from its favorable kinetics. Unlike the auxiliaries discussed in the first chapter, a non-branched α-carbon at the connection site to the peptide gives it higher flexibility and thus improved ligation kinetics and enable ligations at demanding sites. The auxiliary was envisioned to be removed in a radical β-fragmentation mechanism in presence of TCEP with the phenyl ring acting as a radical stabilizer (scheme 2.2 C). The auxiliary's ability to ligate large peptide fragments on solid support was tested by the synthesis of the MUC1 variable number tandem repeat (VNTR) domain. Trials in performing sequential ligation of multiple fragments on solid support were also done.




**Scheme 2.2:** General scheme for auxiliary assisted SPCL of the MUC1 VNTR domain. **A)** N→C directed synthesis using masked thioesters; **B)** C→N directed synthesis using orthogonal thiol protection; **C)** Hypothesized mechanism for auxiliary removal.

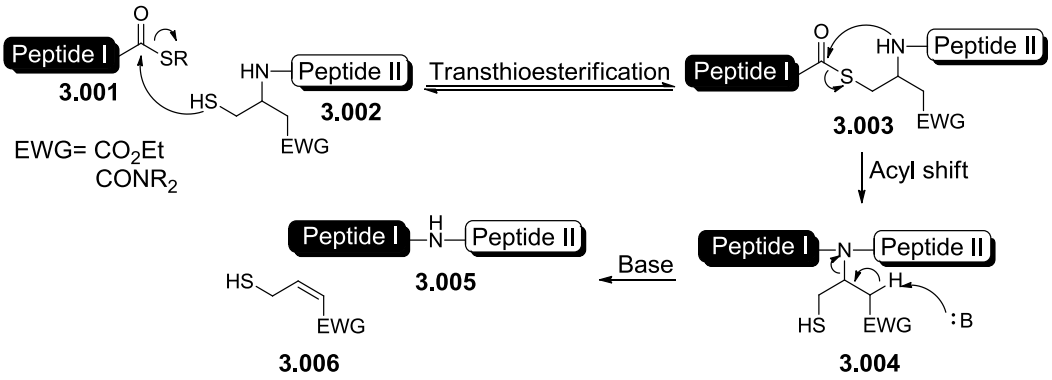
# 3 Results and Discussion

## 3.1 Native Chemical Ligation Assisted by Base-Labile Auxiliaries

The goal of this work is to broaden the scope of auxiliary mediated native chemical ligation. $N^{\alpha}$-auxiliaries facilitate the ligation reaction through a thiol handle which mimics the cysteine function. A specially designed functional group on the auxiliary enables selective cleavage of the auxiliary from the ligation product to obtain the native peptide bond. Most of the existing auxiliaries are acid labile and differ from each other by the sensitivity to acid and to a degree, the scope of ligation sites available by them. This work seeks to develop a new class of auxiliaries, designed to extend the number of available ligation junctions and to be removed not by acidic conditions, which are incompatible with several protein modifications, but rather by exposure to a mildly basic aqueous medium. To fulfill this goal new auxiliaries were designed and synthesized. The auxiliaries possess three functional centers: a thiol moiety for promoting the NCL, a functional group that will enable selective cleavage off the native peptide bond and a third reactive center for coupling to the amino acid or peptide terminal amine. Once the synthesis of the building blocks and auxiliary peptides is completed the auxiliaries' scope and behavior in ligations will be studied on model systems and a method for selective and efficient cleavage of the auxiliary from the ligation product to afford the native peptide will be developed. To prove the robustness and applicability of the method a longer peptide or protein fragment will be synthesized.

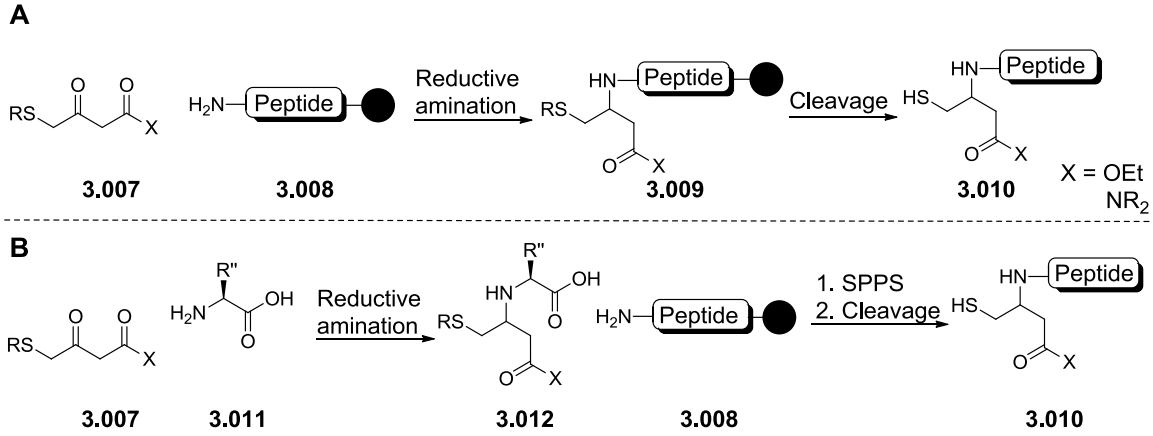
### 3.1.1 Initial Concept and Design of the Auxiliaries

The general scheme and design of the auxiliaries is shown in scheme 3.1. The auxiliaries are based on a butyrate scaffold. The thiol handle is positioned on the terminal carbon, two carbon atoms away from the terminal amine enabling a 5-membered ring transition state (scheme 3.1 **3.003**) as in NCL. The cleavable characteristic of the auxiliary stems from the presence of an electron withdrawing group (EWG). The EWG increases the acidity of the α-proton and was envisioned to facilitate the auxiliary removal under basic conditions by an elimination mechanism (scheme 3.1 **3.004**). The bond between the auxiliary and the peptide backbone is consequentially cleaved to create two stable products: the desired native peptide **3.005** and a α,β-unsaturated butanoic acid derivative **3.006**.




**Scheme 3.1**: General structure and mechanism of base labile auxiliaries.

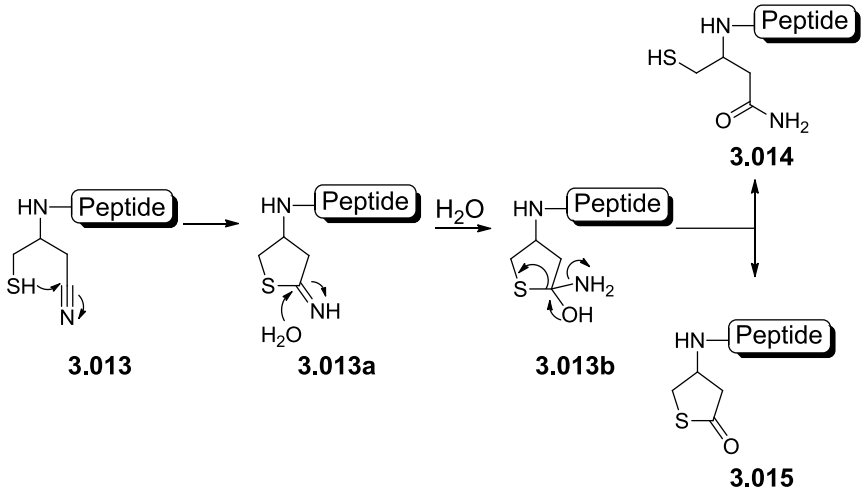
The auxiliaries were designed to be attached directly onto the resin bound peptide through reductive amination between the terminal amine of the peptide and a ketone functional group on the auxiliary precursor. The ability to attach the auxiliary directly to any resin-bound peptide increases the general utility of the method by eliminating the need to synthesize individually modified amino acids for each specific ligation site. If desired however, the auxiliary can be attached to an amino acid to be later used as a monomer for SPPS (scheme 3.2).




**Scheme 3.2**: Strategy for the synthesis of auxiliary modified peptides. **A)** Reductive amination on solid supported; **B)** Coupling of auxiliary modified amino acids.

## 3.1.2 Synthesis of Auxiliary Peptides

Previous results using a cyano moiety as the EWG on the auxiliary have shown that when treated with TFA or ligation buffer, the auxiliary spontaneously rearranges into two new species: amide **3.014** and thiolactone **3.015[**96]. The hypothesized mechanism forming these two species is shown in scheme 3.3. An intramolecular attack on the cyano group by the thiol moiety takes place leading to the formation of imidothioester **3.013a**. This reactive species reacts further in a 1,2-addition of water to form the tetrahedral aminothiophene-alcohol **3.013b** that turns either to the amide species **3.014** or the thiolactone species **3.015** [96].




**Scheme 3.3**: Hypothesized mechanism for the formation of the butanamide **3.014** and thiolactone **3.015** auxiliaries.

As the cyano auxiliary spontaneously formed the amide and ligation experiments done with this auxiliary successfully gave ligation product [96], a study to reach a deeper understanding of the behavior of the amide auxiliary for ligation and removal was planned. Three auxiliaries were designed: The spontaneously formed amide auxiliary **3.017**, the secondary amine N-phenylamide auxiliary **3.018** and the tertiary N-methyl-N-phenylamide **3.019**. The latter two will help in examining the influence substitution of the nitrogen has as well as to see the influence of the electron rich benzene ring on the ligation kinetics and cleavage of the auxiliary. A fourth auxiliary **3.016**, with an ethyl ester as the EWG group, was also designed as it was hypothesized to show similar or improved removal properties due to the ester's greater electronegativity in comparison with an amide.

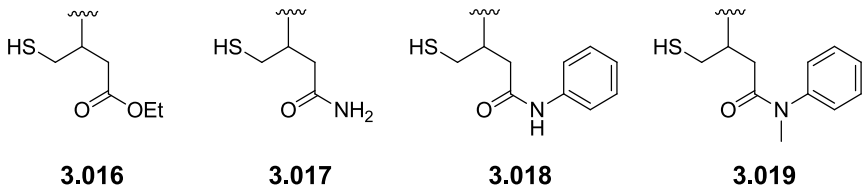
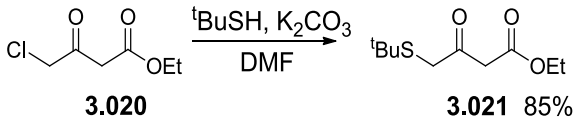



**Figure 3.1**: General structures of auxiliaries.

### 3.1.2.1 Synthesis of Auxiliary Precursors

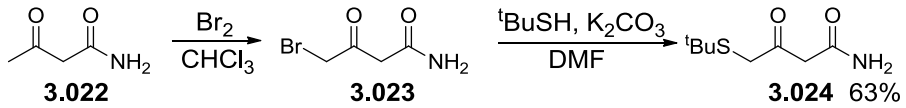
All auxiliary precursors were synthesized bearing a ketone functionality which would later enable the reductive amination between them and the amino acid amine. Auxiliary precursor **3.021** was synthesized in a straightforward manner starting from ethyl 4-chloro-3-oxobutanoate **3.020** by substituting the chloride with tert-butylthiol in the presence of $K_2CO_3$ in DMF. After aqueous work up the crude was passed through a plug of silica to remove any residual impurities to give the pure compound in 85% yield (scheme 3.4).




**3.020** **3.021** 85%

**Scheme 3.4**: Synthesis of auxiliary precursor **3.021**.

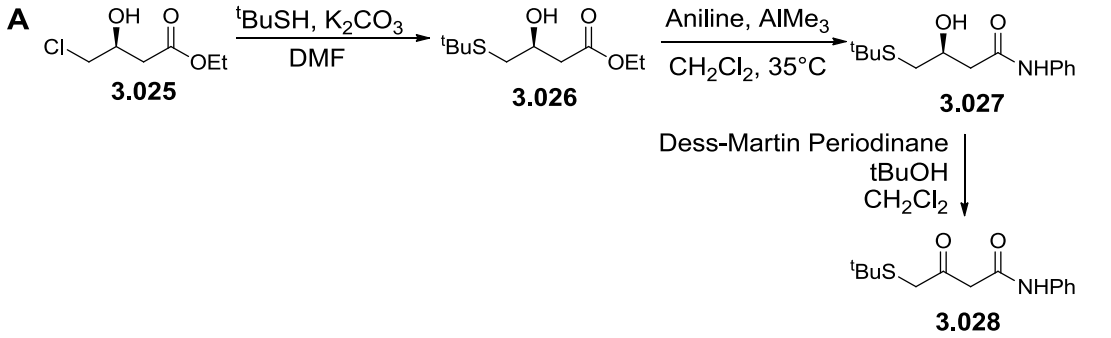
Auxiliary precursor **3.024** was synthesized in two steps. Starting from acetoacetamide **3.022**, bromine was introduced by the unselective bromination- selective debromination reaction [97] yielding 4-bromoacetoacetamide **3.023** which was directly converted into the *S*-tBu protected 4-thioacetoacetamide **3.024** in 63% yield over two steps.

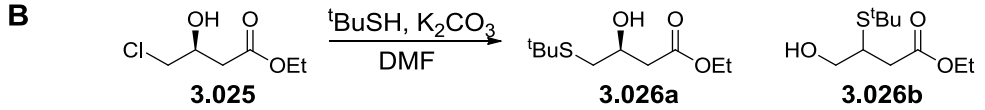



**3.022** **3.023** **3.024** 63%

**Scheme 3.5**: Synthesis of auxiliary precursor **3.024.**

The synthetic route for auxiliary precursor **3.028** is shown in scheme 3.6A. It is a three step synthesis starting from ethyl 4-chloro-3-hydroxybutanoate **3.025**. Attempt to introduce the tert-butylthiol group by employing the conditions used for auxiliary precursor **3.21** led to the formation of two products in nearly 1:1 ratio. NMR analysis revealed that the two products are the regioisomers **3.026a** and **3.026b** (scheme 3.6B).

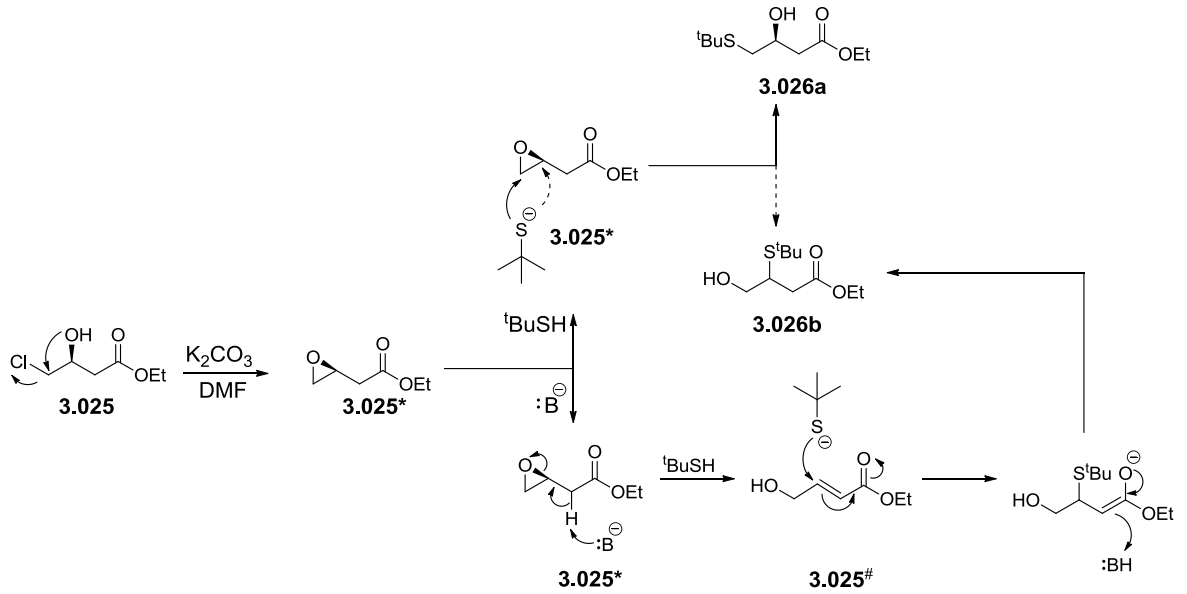



**Scheme 3.6**: **A)** Synthetic route towards auxiliary precursor **3.028**; **B)** Observed products in reaction between **3.025** and tert-butylthiol.

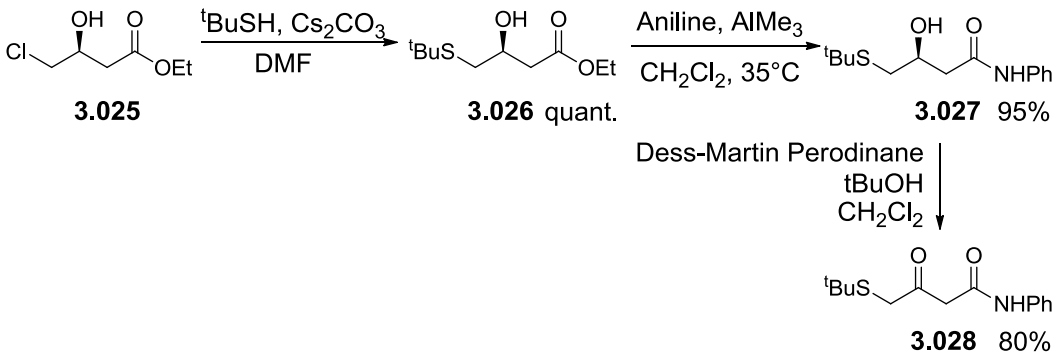
The presence of the two products suggests an initial intramolecular attack of the alcohol on the alkyl chloride forming epoxide **3.025*** (scheme 3.7). A nucleophilic attack by the thiol on the 4$^{th}$ carbon leads to the formation of enantiomer **3.026a** (scheme 3.7). The presence of **3.026b** might be explained by a similar attack on the 3$^{rd}$ carbon, however, this carbon is sterically hindered and an attack on this position is unfavorable. Taking into account the thiol nucleophile itself is bulky as well a different mechanism may play a role in the formation of **3.026b**. A possible route is shown in scheme

3.7. After formation of epoxide **3.025*** the acidic α-proton is abstracted by a base, opening the epoxide and forming the conjugated Michael system ethyl 4-hydroxy-2-butenoate species **3.025**$^{*}$. The Michael system is subsequently attacked by the thiol in a 1,4-addition mechanism to give product **3.026b**.



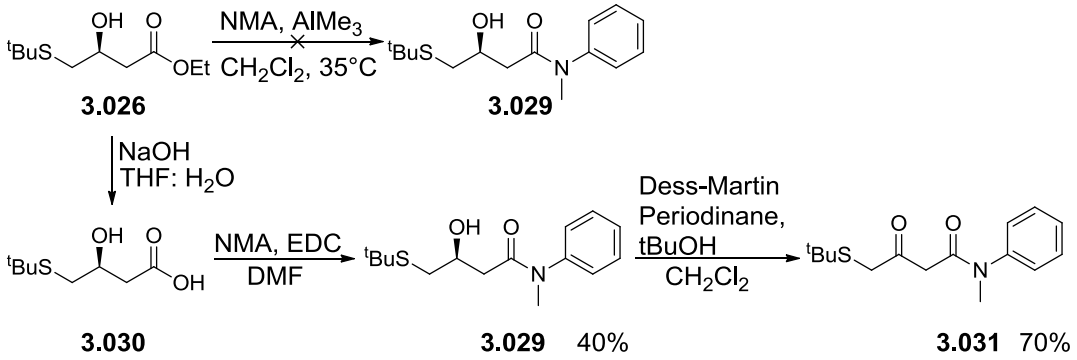

**Scheme 3.7**: Hypothesized mechanisms for generation of products **3.026a** and **3.026B** through epoxide opening.

To overcome this problem the reaction conditions were slightly altered. First, more equivalents of base and thiol were used. In addition, the stronger base[98], cesium carbonate, was used. As an added measure the thiol and base were allowed to react for several minutes to pre-form the thiolate species before adding chloride **3.025**. Under these conditions, a clean conversion to the desired product was achieved in quantitative yield. The crude was converted directly into the N-phenylamide **3.026** through trimethylaluminum amidation of esters [99]. The final step was the formation of the ketone. To perform this reaction a mild oxidation reaction that would not oxidize the thioether was required. The Dess-Martin periodinane reagent [100] proved to be suitable, leading to a fast and clean conversion with excellent yield (scheme 3.8).




**Scheme 3.8**: Synthesis of auxiliary precursor **3.028**.

The final precursor **3.031** was synthesized in a similar fashion to auxiliary **3.028**. However, a direct amidation of the ester was not possible with the secondary, more sterically demanding amine of the N-methylaniline (NMA). To this end, ester **3.026** was first hydrolyzed with NaOH in THF and the amine was then coupled to the acid using EDC in DMF to give the amide **3.029** with 40% yield. Dess-Martin oxidation gave the desired ketoamide compound **3.031** in 70% yield (scheme 3.9).




**Scheme 3.9**: Synthesis of auxiliary precursor **3.031**.

With all four auxiliary precursors in hand, the next step, attachment of the auxiliaries on to the model peptides by reductive amination, could begin.

### 3.1.2.2 Reductive Amination on Solid Support

The model peptidyl resins GRAEYSGLG **3.032** and ARAEYSGLG **3.033** were prepared employing routine Fmoc SPPS protocol on a Rink Amide resin (figure. 3.2)

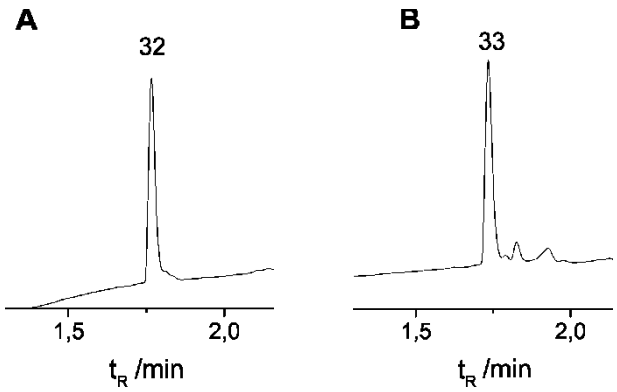
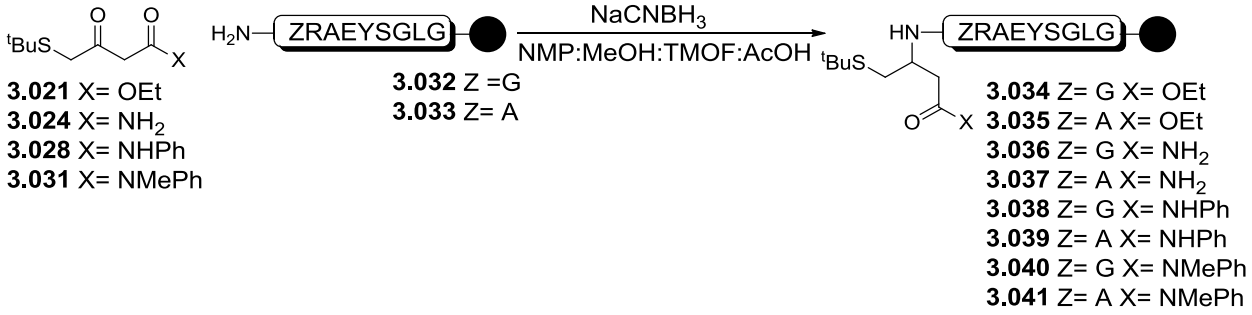



**Figure 3.2**: UPLC analysis of crude model peptides **A)** GRAEYSGLG **3.032**; **B)** ARAEYSGLG **3.033**; 03-50% B in 20 min.

The auxiliary precursors were coupled to the resin bound peptides by reductive amination. The auxiliary precursors were dissolved in 1:1:1 NMP:MeOH:TMOF and 10% v/v acetic acid and added to the resin that was pre-swelled in NMP. After 60 min, sodium cyanoborohydride in 1:1 NMP:MeOH

was added and the progress of reaction was monitored by HPLC-MS. The analysis showed product was forming, however, a major side product was also emerging. Analysis of the reaction revealed N-acetylation. This unwanted reaction was probably due to a reaction of the amine with the acetic acid. To avoid it, the auxiliary precursors in same solvent mixture, were added to the resin and left to react for 5 min before adding 5% v/v of acetic acid. Postponing the addition of the acetic acid together with halving its volume helped to reduce the amount of N-acetylation drastically. Comparing the efficiency of the reductive amination between the four precursors, no startling differences were observed. The N-methyl-N-phenylamide auxiliary precursor **3.031** has exhibited slightly less favorable reaction profile, requiring longer reaction time to reach a satisfactory conversion ratio. However, a major difference was seen when reaction was performed on glycine in comparison to alanine. While on the N-terminal glycine peptide **3.032** reaction reached 80-90% conversion after 24-48 hr using 10-15 eq. of auxiliary precursors, the N-terminal alanine peptide **3.033** required the use of 20-30 eq. and often 3-5 days to achieve acceptable conversion to product. In the case of the N-methyl-N-phenylamide precursor **3.031**, the reaction did not reach more than 60-70% conversion even when allowed to react for extended period of times and with fresh batch of auxiliary precursor added. Once reaction was complete or, alternatively, did not progress further, the peptide was cleaved from support with a cleavage mixture of TFA:$H_2O$:TIS 95:2.5:2.5 v/v to yield the crude *S*-tBu protected auxiliary peptides (figure 3.3)



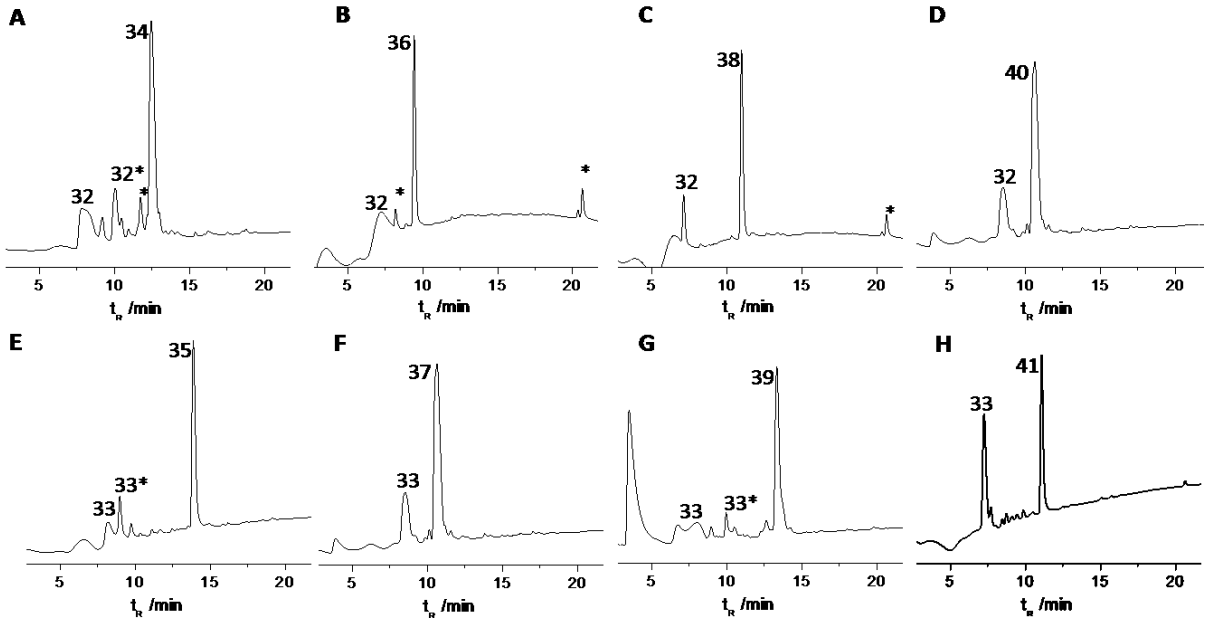
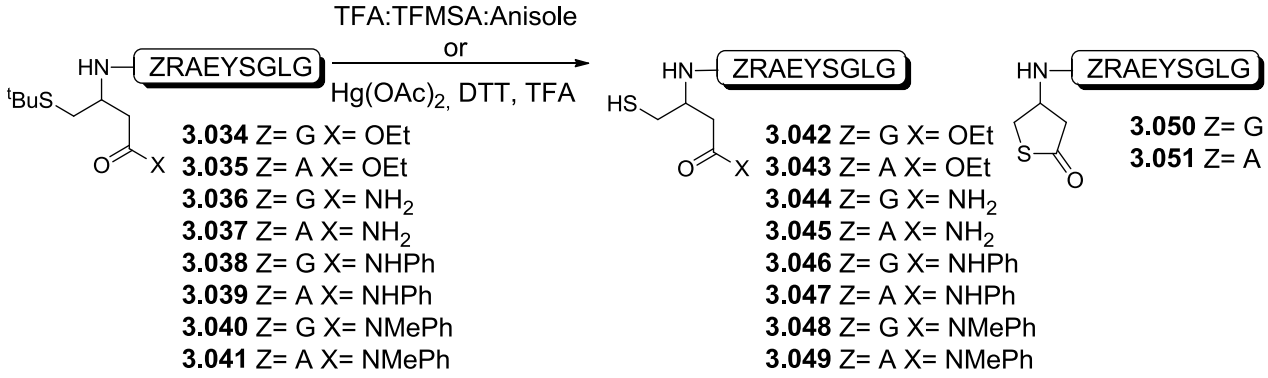



**Figure 3.3**: HPLC analysis of reductive amination on resin bound peptides **3.032** and **3.033** to give crude N-terminal glycine auxiliary peptides **3.034 (A), 3.036 (B), 3.038 (C), 3.040 (D) a**nd N-terminal alanine auxiliary peptides **3.035 (B), 3.037 (F), 3.039 (G),3.041 (H)**; Peak **3.032*** and **3.033*** corresponds to N-acetylation of peptide **3.032** and **3.033** respectively; peak labeled with * corresponds to unassigned masses; 03-60% B in 20min, 210 nm.

### 3.1.2.3 Deprotection of the tert-Butyl Protection Group

To remove the tert-butyl group, peptides were first treated with a cold mixture of TFA:TFMSA:anisole at 0 °C for 90 min. Under these conditions the auxiliary peptides were divided into two groups based on their response to the cleavage mixture. While the N-phenylamide and N-methyl-N-phenylamide auxiliary peptides **3.038-3.041** yielded the desired fully unprotected auxiliary peptides **3.046-3.049** respectively (figure 3.4 C,D,G,H) as the main products, the ethyl ester and amide auxiliary peptides **3.034-3.037** formed almost exclusively the thiolactone modified peptides **3.050** and **3.051** respectively (see p. 116). In an effort to hinder this side reaction, a different procedure for removal of the tert-butyl group from peptides **3.034-3.037** was used, employing mercury acetate [101]. Under mercury acetate treatment the thiolactone formation was completely inhibited or greatly reduced (figure. 3.4 A,B,E,F).



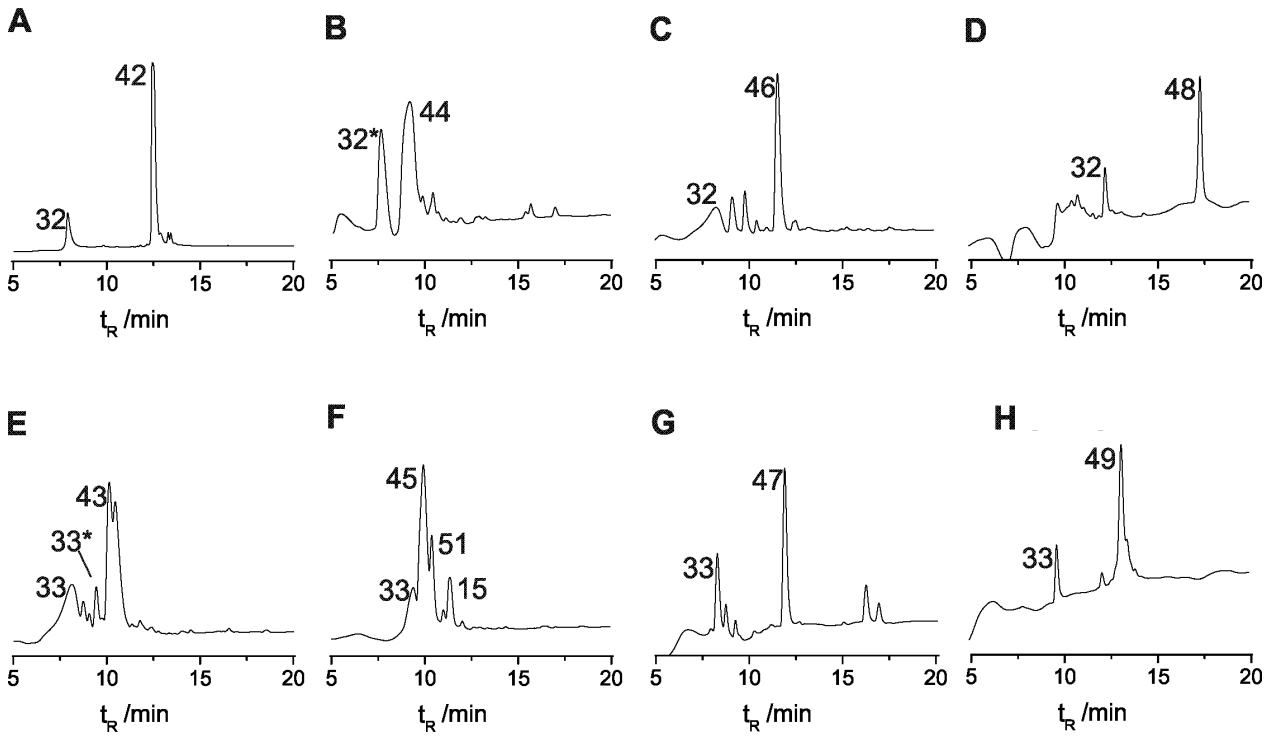
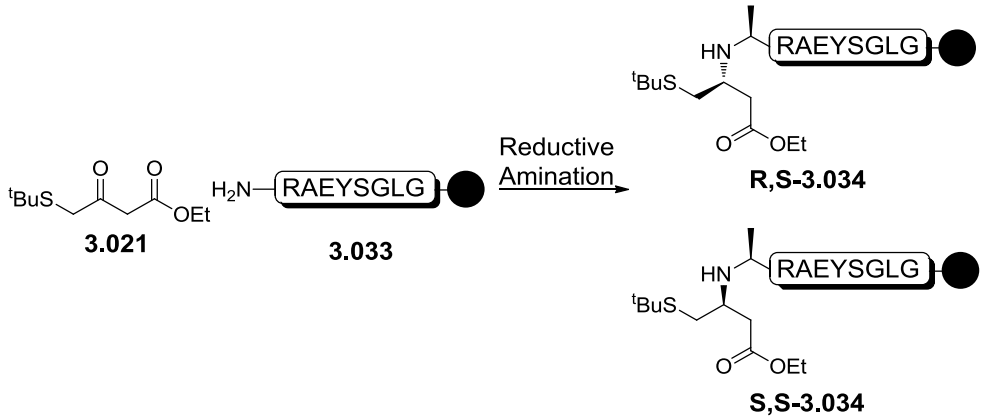



**Figure 3.4**: HPLC analysis of tert-butyl deprotection to give unprotected auxiliary peptides. A) reaction of peptide **3.034** to give **3.042**; B) reaction of peptide **3.036** to give **3.044**; C) reaction of peptide **3.038** to give **3.046**; D) reaction of peptide **3.040** to give **3.048**; E) reaction of peptide **3.035** to give **3.043**; F) reaction of peptide **3.037** to give **3.045**; G) reaction of peptide **3.039** to give **3.047**; H) reaction of peptide **3.041** to give **3.049**; peaks **3.032** and **3.033** correspond to peptides GRAYESGLG and ARAEYSGLG respectively; peak **3.032*** and **3.033*** correspond to N-acetylation of peptide **3.032** and **3.033** respectively; Gradient A-C, E,F,H: 03-60% B in 20min, Gradient D, G: 03-40% B in 20 min, 210 nm.

As seen in figure 3.4 the removal of the thiol protecting group from ethyl ester auxiliary peptides **3.034-3.035**, the N-terminal glycine amide auxiliary peptide **3.036** and N-phenylamide and N-methyl-N-phenylamide auxiliary peptides **3.038-3.041** gave a direct conversion from starting material to product with impurities stemming from incomplete reductive amination of peptides GRAEYSGLG **3.032** or ARAEYSGLG **3.033** as well as the respective N-acetylated derivatives **3.032*** and **3.033***. The removal of tert-butyl group from the N-terminal alanine amide auxiliary peptide **3.037** resulted in formation of thiolactone peptide **3.051** (figure 3.4 F) demonstrating the instability of the amide auxiliary.

Figure 3.4 E shows the HPLC analysis of crude material obtained after *S*-tBu removal from ethyl ester auxiliary peptide **3.035**. In this trace, a double peak for product **3.043** is observed. Mass analysis showed the two peaks to have identical mass corresponding to the desired product **3.043**. The double peak probably is the result of two diastereomers formed during reductive amination. As the reductive amination is non-stereospecific both possible diastereomers will be formed (scheme 3.10). Interestingly, the differentiation into two peaks is not always observed.




**Scheme 3.10**: Evolvement of R,S- and S,S-enantiomers during reductive amination.

The auxiliary peptides were purified on a preparative HPLC to yield the pure auxiliary peptides **3.042-3.049** ready for ligation experiments. Yields and mass analysis for each of the auxiliary peptides is shown in table 1.

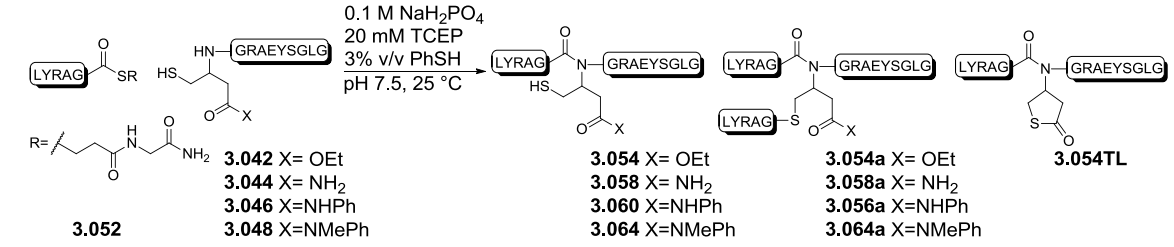
**Table 1: Mass analysis and isolated yields for tert-butyl removal from peptides 3.034-3.041 to give peptide auxiliaries 3.042-3.049**

| Identity | Entry | Peak | Observed Mass (Da) | Calculated Mass (Da) | Isolated Yield |
|---|---|---|---|---|---|
| **$CO_2Et$ (G)** | A | 3.032 | 907.7 | 908.0 | |
| | | 3.042 | 1053.8 | 1054.2 | 30% |
| **$CONH_2$ (G)** | B | 3.032 | 907.7 | 908.0 | |
| | | 3.032* | 949.8 | 950.0 | 24% |
| | | 3.044 | 1024.8 | 1025.2 | |
| **CONHPh (G)** | C | 3.032 | 907.7 | 908.0 | |
| | | 3.032* | 949.8 | 950.0 | |
| | | 3.050 | 1007.7 | 1008.2 | 36% |
| | | 3.046 | 1100.9 | 1101.3 | |
| **CONMePh (G)** | D | 3.032* | 949.8 | 950.8 | |
| | | 3.048 | 1114.9 | 1115.9 | 16% |
| | | 3.040 | 1171.0 | 1171.4 | |
| **$CO_2Et$ (A)** | E | 3.033 | 921.7 | 922.0 | |
| | | 3.043 | 1067.8 | 1068.2 | 14% |
| **$CONH_2$ (A)** | F | 3.033 | 921.7 | 922.0 | |
| | | 3.045 | 1038.6 | 1039.2 | |
| | | 3.051 | 1021.6 | 1022.1 | |
| | | 3.037 | 1094.9 | 1095.3 | 18% |
| **CONHPh (A)** | G | 3.033 | 921.7 | 922.0 | |
| | | 3.033* | 963.7 | 964.0 | |
| | | 3.051 | 1021.8 | 1022.1 | 15% |
| | | 3.047 | 1114.8 | 1115.3 | |
| **CONMePh (A)** | H | 3.033 | 921.8 | 922.0 | |
| | | 3.049 | 1128.8 | 1129.3 | 15% |

## 3.1.3 Ligation Experiments

With the auxiliary peptides in hand ligation studies begun,. Ligations were performed at a concentration of 5 mM in 100 mM phosphate buffer pH 7.5 at 25 °C with 20 mM TCEP and 3% thiophenol as additives. Reactions were monitored using UPLC-MS.

### 3.1.3.1 Ligations at Glycine-Glycine Junction



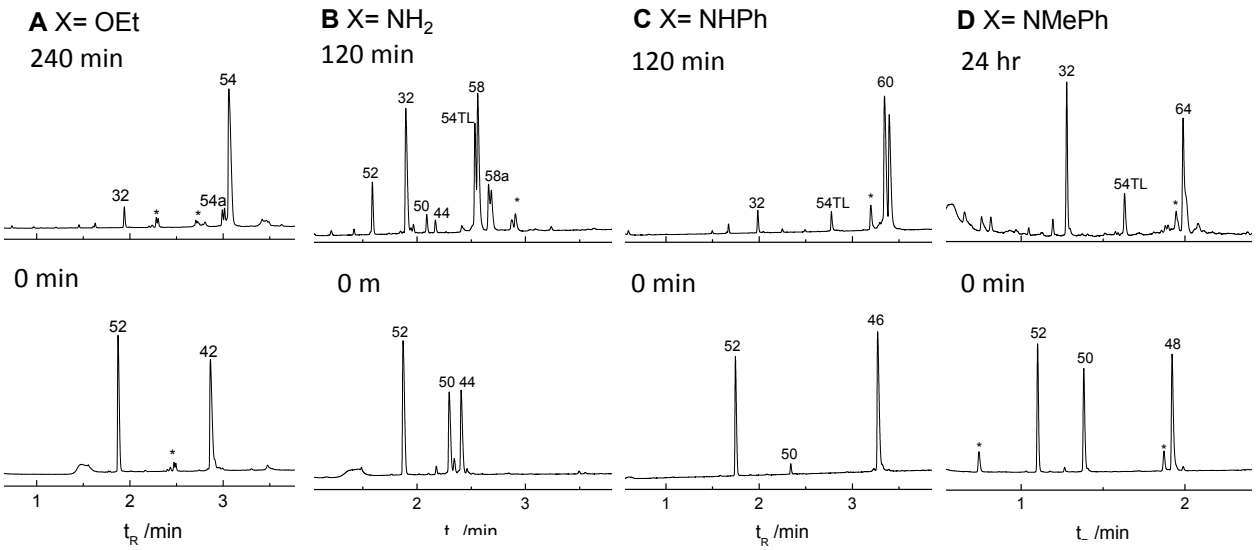




**Figure 3.5**: UPLC analysis of ligation between LYRAG thioester **3.052** and auxiliary peptides **3.042 (A)**, **3.044 (B)**, **3.046** (**C**) and **3.048(D)**. **3.032** GRAEYSGLG, **3.050**: Thiolactone-GRAEYSGLG, peaks labeled with ***** correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

Figure 3.5 shows the UPLC analyses of ligations between the four auxiliary peptides **3.042**, **3.044**, **3.046** and **3.048** with thioester **3.052**. Two product classes can be observed: ligation product and derivatives of it (double N- and S-acetylated ligation product marked Xa and thiolactone species marked XTL) and peptide GRAYESGLG **3.032**. These products are formed as a result of two competing reactions: auxiliary mediated NCL and removal of auxiliary from auxiliary peptides. The factors determining the outcome of the ligation are therefore rate of ligation vs. rate of auxiliary cleavage to form the native unmodified peptide **3.032**. Plotting the time vs. yield of the four reactions and estimating the initial rate constants shows that except the distinctly slower N-methyl-N-phenylamide auxiliary peptide **3.048**, the difference in initial reaction rates between ethyl ester **3.042**, amide **3.044** and N-phenylamide **3.046** auxiliary peptides are negligible, ranging from 0.03-0.04 $M^{-1}s^{-}$

[1] (figure 3.6 A). This means that it is the rate of the auxiliary removal that governs the achievable yields of auxiliary modified ligation product. By examining the UPLC traces for reactions systems analyzed immediately after reagent mixing (t = 0 min in fig. 3.5) one sees that ethyl ester **3.042** and N-phenylamide **3.046** auxiliary peptides are stable (figure 3.5 A,C), while amide **3.044** and N-methyl-N-phenylamide **3.048** auxiliary peptides give a 1:1 mixture of thiolactone peptide **3.050** and auxiliary peptide (figure 3.5 B,D). A closer inspection of the ligation progress hinted that the thiolactone species is being formed over time. Once formed it does not participate in ligation but instead rapidly decomposes to give unmodified peptide **3.032**. To prove this, ligation with thiolactone peptide **3.050** was conducted, clearly demonstrating its removal to yield peptide **3.032** as a single product (figure 3.6 B). The propensity to form the thiolactone is especially prominent with the amide and the N-methyl-N-phenylamide auxiliaries **3.044** and **3.046** but is not exclusive to them. Thiolactone is also seen with the N-phenylamide auxiliary **3.045** albeit in low concentrations. Thiolactonisation also occurs on the ligation product (figure3.5 B-D **54TL**). Ligation products **3.054** and **3.060** were purified by semi-preparative HPLC with isolated yield of 56% each. The ligation products **3.058** and **3.064** were not isolated due to the poor behavior exhibited by these auxiliaries making them unsuitable for auxiliary assisted ligation.

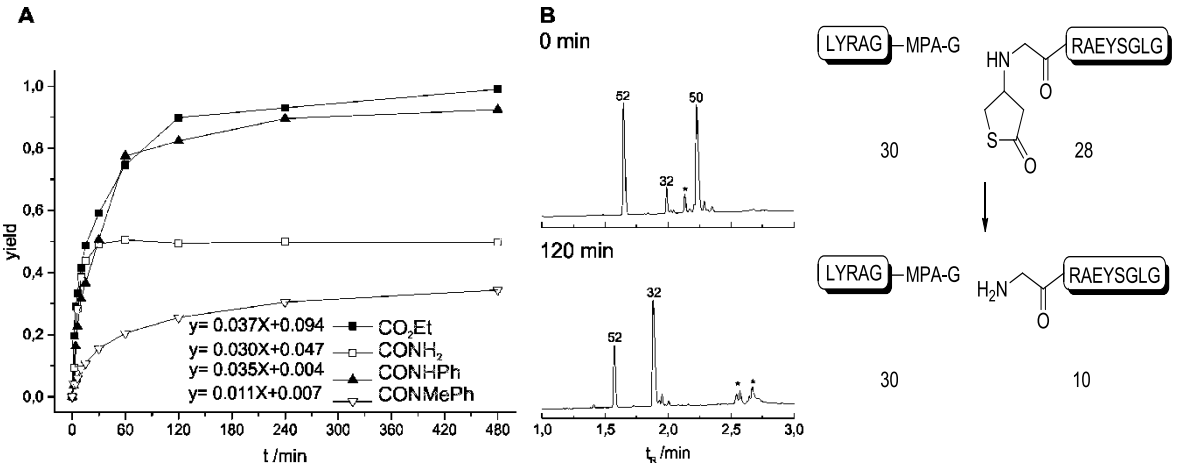



**Figure 3.6**: **A)** time vs. yield and initial rate constants of ligation between thioester **3.052** and auxiliary peptides **3.042**, **3.044**, **3.046**, **3.048**; **B)** UPLC analysis of reaction between thioester **3.052** and thiolactone peptide **3.050** at 0 hr and 2 hr; **C)** Scheme of products formed during reaction between peptide thioester **3.052** and thiolactone peptide **3.050**; peaks labeled with * correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

### 3.1.3.2 Ligation at Alanine-Glycine Junction

Experiments at Ala-Gly junction further emphasized the behavior observed at the Gly-Gly junction. The increased sterical demand of this junction lowered the rate of ligation significantly increasing the formation of GRAEYSGLG peptide **3.032** in all cases with conversion of ~10% for ethyl ester and N-phenylamide auxiliary peptides **3.042** and **3.046** and ~35% for amide and N-methyl-N-phenylamide peptides **3.044** and **3.048** according to integration (figure 3.7). Ethyl ester auxiliary peptide **3.042** and N-phenylamide auxiliary peptide **3.046** showed 65% and 80% conversion to ligation

product respectively after 4 hr and isolated yields of 40% for ethyl ester auxiliary **3.055** and 37% for N-phenylamide auxiliary **3.061** (figure 3.8 A and C). Ligation with peptides **3.044** and **3.048**, however, became impractical due to their tendency to rapidly form the thiolactone derivative **3.050** and less than 30% conversion to ligation product observed in both reactions (figure 3.8 B and D).

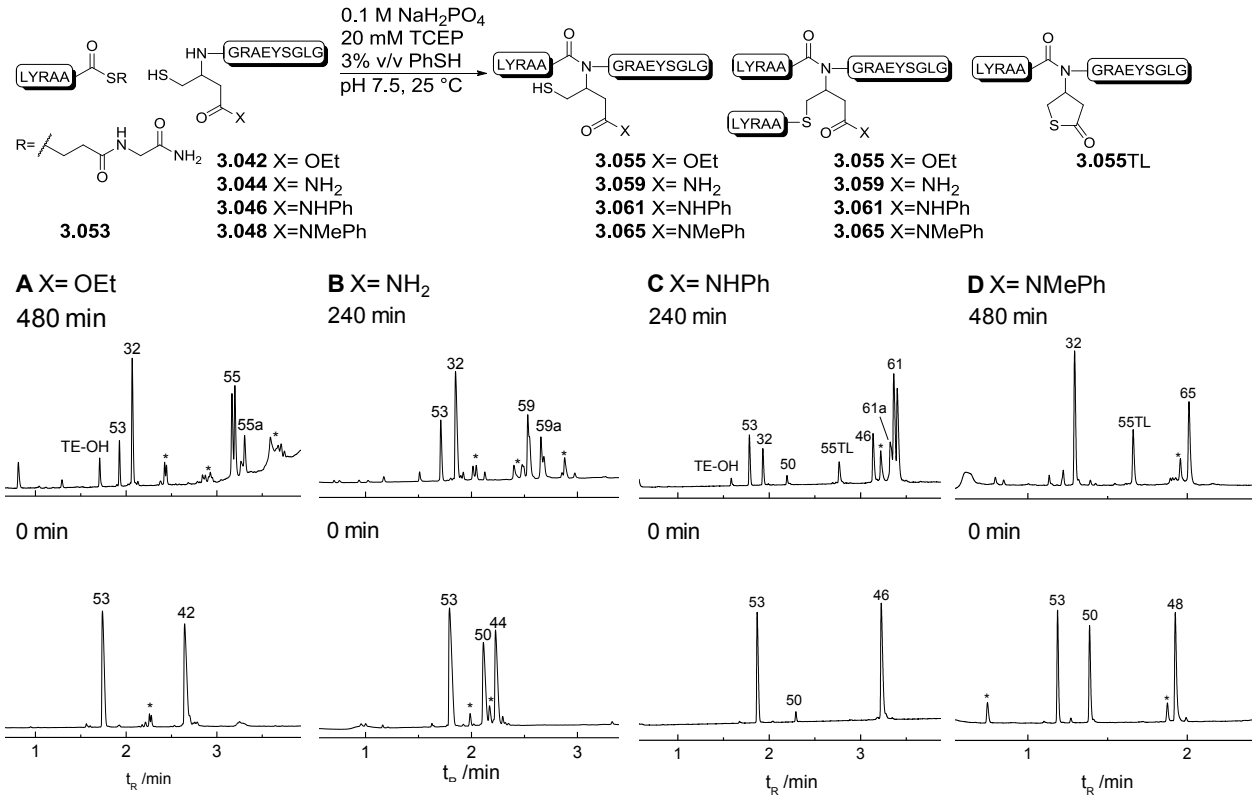



**Figure 3.7**: UPLC analysis of ligation between LYRAA thioester **3.053** and auxiliary peptide **3.042** (**A**), **3.044** (**B**), **3.046** (**C**), and **3.048** (**D**); **3.032**: GRAEYSGLG, **3.050**: Thiolactone GRAEYSGLG, **TE-OH**: LYRAA-OH; peak labeled with ***** correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

Calculation of the initial reaction rates demonstrated the strong influence the added sterics had on ligation with an astonishing 2 orders of magnitude for N-methyl-N-phenylamide auxiliary peptide **3.048** (figure 3.8). With the poor results achieved with amide auxiliary peptide **3.044** and N-methyl-N-phenylamide auxiliary peptide **3.048**, due to their tendency to convert to the inactive thiolactone species, they have been discarded from further experimentation.

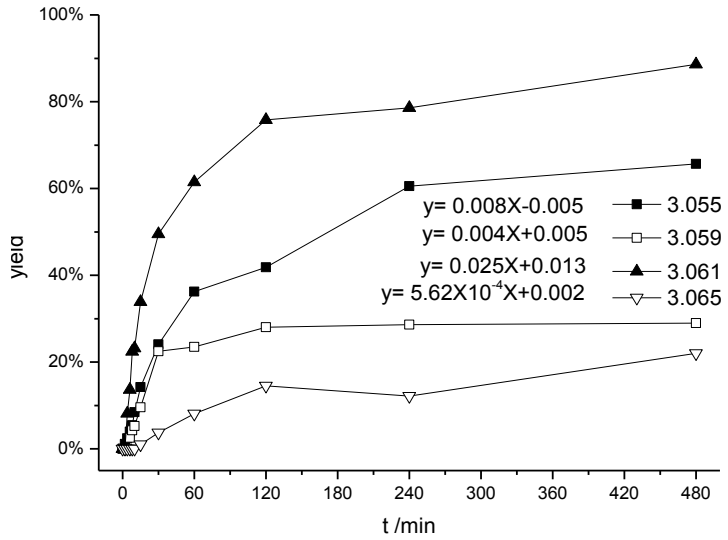



**Figure 3.8**: time vs. yield plot and initial rate constants of ligation between LYRAA thioester **3.053** and auxiliary peptides **3.042**, **3.044**, **3.046**, **3.048** to give ligation products **3.055, 3.059, 3.061 and 3.065** repectively

### 3.1.3.3 Ligation at Glycine-Alanine and Alanine-Alanine Junctions

Comparing the Gly-Ala junction to its "mirror image", the Ala-Gly junction (figure 3.9), shows that the ethyl ester auxiliary **3.016** and the N-phenylamide auxiliary **3.018** are still able to promote ligation between the two peptide fragments. However, the added sterical hindrance on the auxiliary peptide had a greater effect on rate and yield of the ligation than in the previous case, when the increased sterical demand was located on the thioester fragment. The rate is halved for the ethyl ester auxiliary peptide **3.043** and is more than 80% slower for the N-phenylamide auxiliary **3.047**. The conversion to ligation product is 50% of the value measured for the Ala-Gly junction and isolated yields were 10% for ligation product **3.056** ($CO_2Et$) and 15% for ligation product **3.062** (CONHPh).

With the Ala-Ala junction the scope of the auxiliaries is reached. In this junction, the sterical demand is too high to enable the S→N acyl shift and no product formation occurs (figure 3.10). Examination of the ligation progress over time reveals the formation of transthioesterification intermediate followed by hydrolysis to yield the hydrolyzed thioester and the auxiliary peptide. The latter disassemble via the thiolactone to give peptide ARAEYSGLG **3.033** as already observed before (figure 3.9).

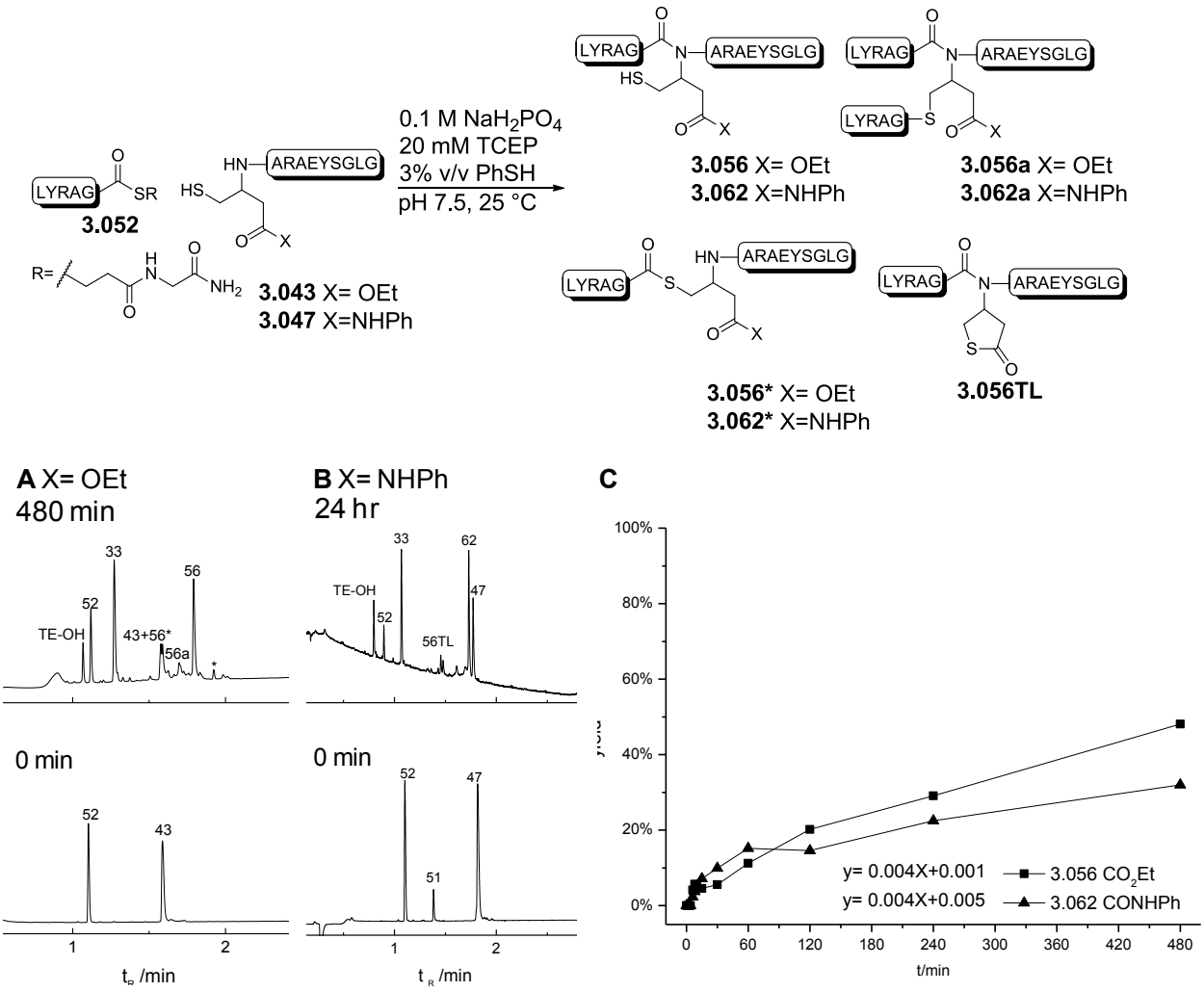



**Figure 3.9**: UPLC analysis of ligation between LYRAG thioester **3.052** and auxiliary peptides **3.043** (**A**) and **3.047** (**B**); **C)** Time vs. yield and initial rate constants for the ligation **A B**; **3.033**: ARAEYSGLG, **3.051**: Thiolactone ARAEYSGLG (TL), **TE-OH**: LYRAG-OH; peak labeled with ***** correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

The inability of the auxiliaries to mediate the ligation at the Ala-Ala junction can be explained by the kinetic profile of auxiliary assisted NCL. Unlike NCL, where transthioesterification is the rate determining step, auxiliary assisted NCL rate is dictated by the S→N acyl shift as indicated by the detection of the thioester intermediate. Analyzing the ligation, one sees a rapid disappearance of starting materials and accumulation of the thioester intermediate **3.063*** which consequently rearranges to give the ligation product (Figure 3.10 A). Due to the auxiliary added sterical hindrance on the amine, the S→N acyl shift becomes very sluggish or completely blocked. Under these conditions only the removal of the auxiliary takes place leading to peptide **3.033** being the major product of the reaction.

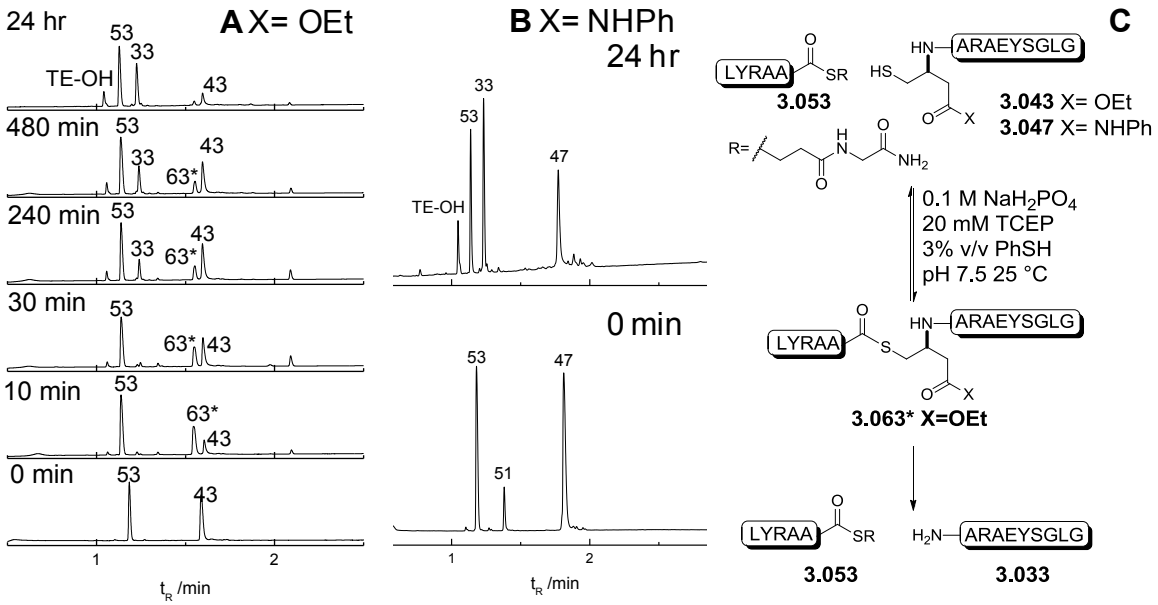



**Figure 3.10**: UPLC analysis of ligation between LYRAA thioester **3.053** and auxiliary peptides **3.043** (**A**) and **3.047** (**B**); **C)** Scheme of reaction between auxiliary peptide **3.043** or **3.047** with thioester **3.053**; **3.051**: Thiolactone ARAEYSGLG (TL), **TE-OH**: LYRAG-OH, 03-60% B in 6 min, 210 nm.

### 3.1.3.4 Premature Removal of Auxiliary

In all ligation studies the removal of the auxiliary from the auxiliary peptide was observed. According to the working hypothesis that was the basis of the initial design, the auxiliary should have remained stable on the amine, only turning labile once an amide bond is formed. The principle behind this assumption was that the amine, being a base, is a very poor leaving group while the poorly basic amide makes for a good leaving group. This characteristic was the working principle for previous $N^{\alpha}$-auxiliaries. In the cases studied here, however, an opposite tendency was observed: the auxiliaries were removed from the unreacted auxiliary peptides but once the amide bond has been formed the auxiliaries' stability seemed to increase (as no auxiliary-free ligation product was detected).

The formation of the thiolactone was especially prominent for amide auxiliary **3.017** and N-methyl-N-phenylamide auxiliary **3.019** which undergo cyclization even at storage as seen in the time zero chromatograms where roughly a 1:1 ratio between the linear and cyclic auxiliary peptides was observed (figure 3.5 B and D and figure 3.11).

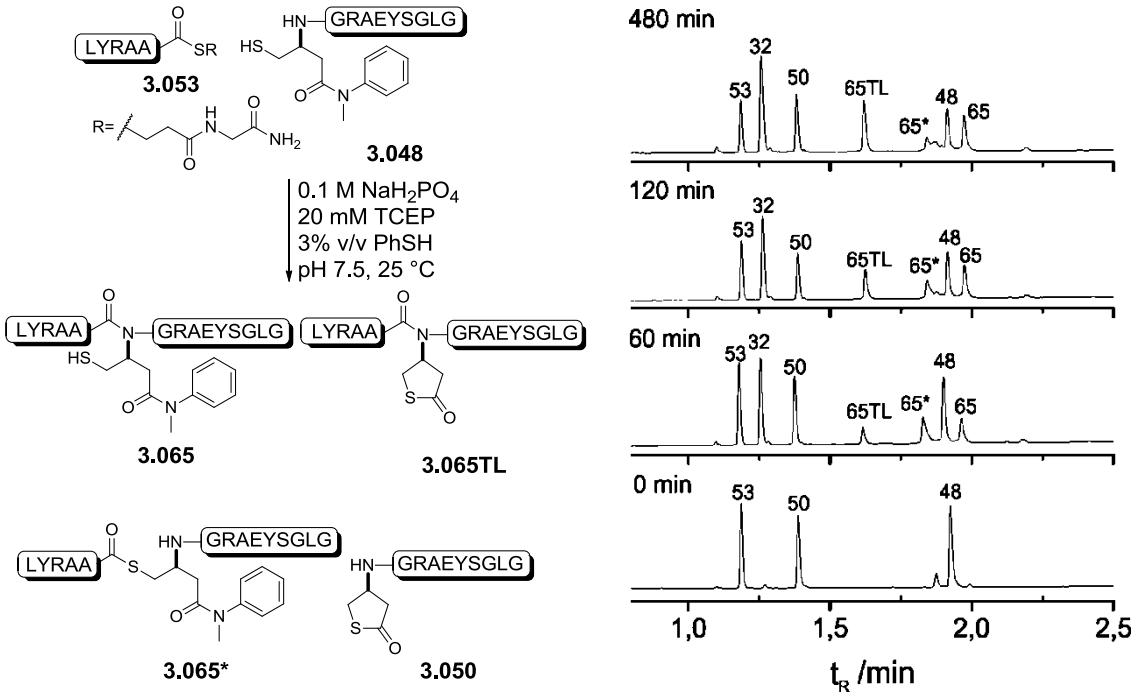
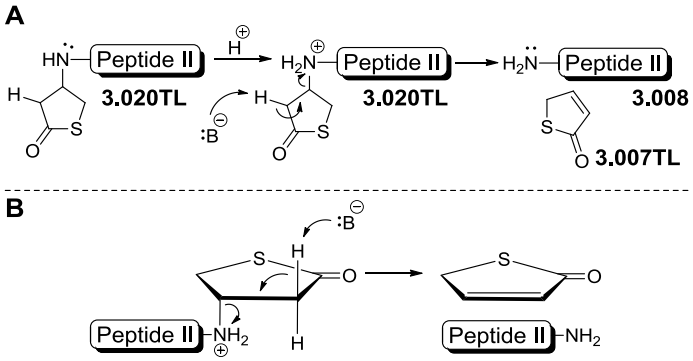



**Figure 3.11**: Scheme and UPLC analysis of ligation between thioester **3.053** and auxiliary peptides **3.048**. 03-60% B in 6 min, 210 nm.

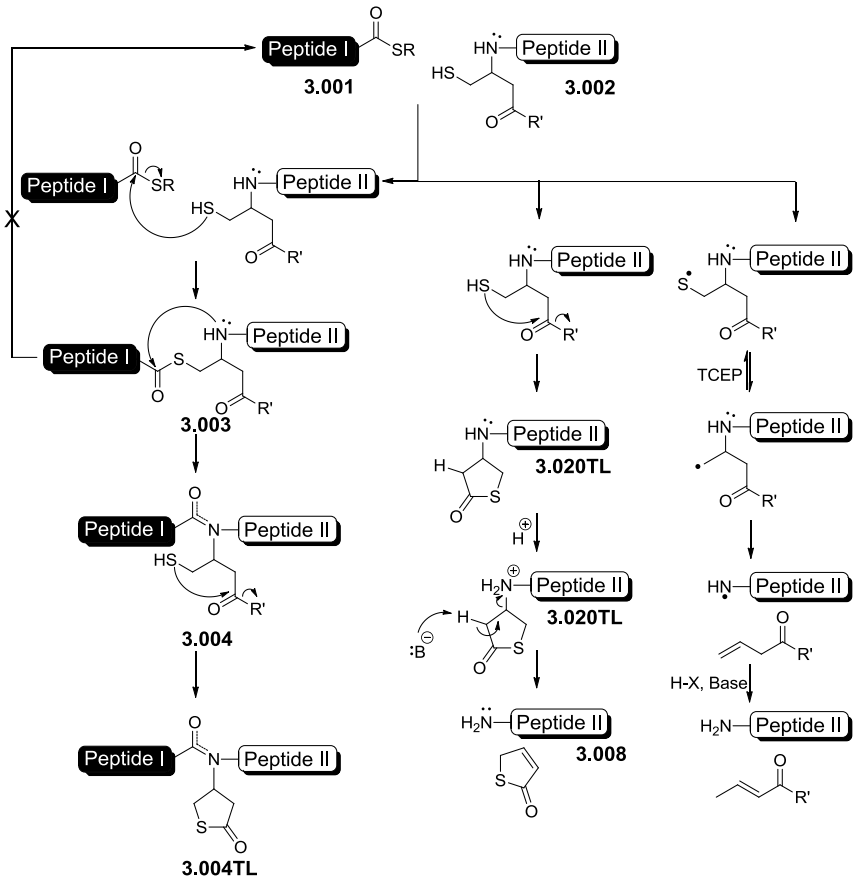
The observation that the stability of the auxiliaries is higher once an amide bond has been formed may be the result of amine protonation forming the ammonium salt,. The much less basic amide is not protonated under the ligation conditions resulting in the observed stability of the auxiliary once ligation product was formed (scheme 3.11 A). This pathway also provides an explanation to the cyclic auxiliary higher elimination rate as it locks the amine and the α-proton in an anti configuration required for an E2 mechanism (scheme 3.11 B).




**Scheme 3.11**:.A) Premature cleavage of the auxiliary triggered by protonation of the amine to from thr ammonium salt; B) Thiolactone formation locks the ammonium and hydrogen in required anti-conformation for E2 elimination.

As long as the sterical demand at the ligation junction was low and formation of ligation product rapid, the presence of these species remained negligible. However, even a slight increase in sterical demand led to major formation of peptides **3.032** or **3.033** with the effect becoming more prominent when the added sterical hindrance was located on the auxiliary peptide.

In these cases a thioester intermediate was detected (Figure 3.10 A peak 63*) but failed to rearrange to give the native peptide bond fast. It consequently dissociated back to the two peptide fragments either by hydrolysis or thiolysis. The auxiliary peptide reacted intramolecular to form the highly favorable 5-memberd ring thiolactone that cannot participate in ligation and decomposes to give the unmodified peptides. A hypothesized pathway for formation of all seen products is shown in scheme 3.12.




**Scheme 3.12**: Hypothesized pathway of auxiliary-assisted NCL yielding all observed species.

### 3.1.3.5 Ligation Experiments Summary

The ligation studies have shown all four auxiliaries to share a similar behavior. As long as the steric limitation remained low the ligations proceeded well to yield product with high conversion and good isolated yields. The auxiliaries were capable to promote ligation as long as the ligation junction contained glycine residues albeit with a drastic decrease in rate and product formation in more sterically demanding sites. The limit of the auxiliary ability to promote ligation was met when Ala-Ala junction was tried. In this case the S→N acyl shift is not possible due to the combined steric hindrance of the auxiliary backbone and the alanine residues methyl groups. In this regard, the auxiliaries discussed here exhibit similar reactivity as previously published auxiliaries [102].

Despite a design concept based on a linear butyrate scaffold aimed to give the auxiliaries flexibility and thus extend the scope of junctions that can be ligated with their assistance, the

auxiliaries failed to promote ligation at sterically hindered junctions. The suggested explanation for this is the different kinetic profile of NCL and auxiliary-assisted NCL. In NCL the rate determining step is the transthioesterification while in auxiliary assisted ligations it is the thioester rearrangement to give the native peptide bond that is rate limiting. This kinetic characteristic of auxiliary assisted ligations is well known[102], and is most likely due to the secondary amine increased sterical hindrance. The auxiliaries presented here all share a common scaffold with a tertiary carbon attached to the terminal amine. As shown in NCL[23] and other published works extending NCL such as ligation at valine[36], ligations are highly susceptible to sterical hindrances slowing the reaction drastically or preventing the reaction to occur altogether. Here, the bulkiness of the auxiliaries prevents the S→N acyl shift when the sterical demand of the ligation site increases. Peptide auxiliaries with a non-branched structure adjacent to the amine have proven to be highly effective in ligation of sterically hindered junctions [95].

The auxiliaries have a characteristic that hampers their usability: the auxiliaries' lability under ligation conditions. A peak corresponding to GRAEYSGLG peptide **3.032** or ARAEYSGLG peptide **3.033** is observed with varying abundance in all ligations performed. These peptides were formed due to the removal of the auxiliary seemingly in a two-step fashion: a slow cyclization step gives the thiolactone species, followed by elimination of the thiolactone. The cyclization reaction differs between the auxiliaries. The tendency to cyclize is highest for the amide auxiliary **3.017**, followed by the N-methyl-N-phenylamide **3.019**, the N-phenylamide auxiliary **3.018** and last the ethyl ester auxiliary **3.016**. While not proven, a removal directly from the linear form probably takes place as well. Interestingly, under ligation conditions the auxiliaries remain stable once ligation product was formed.

As the auxiliary removal is a competing reaction to the ligation, the factor determining which path becomes prominent, ligation product formation or removal of auxiliary to give peptide **3.032** or **3.033**, is determined by the rate difference between the S→N acyl shift and the cyclization to thiolactone. As the ligation junction hindrance increases the acyl rearrangement rate decreases making the auxiliary removal route the dominant path.

Despite the auxiliaries' lability and inability to excess sterically demanding junctions, the ethyl ester auxiliary **3.021** and the N-phenylamide auxiliary **3.028** showed most promise, exhibiting higher stability, faster reaction rates and less side product formation. Auxiliaries **3.024** and **3.031** were discarded from further investigation due to their instability. Both auxiliary **3.021** and **3.028** showed good results during the ligation studies with highest rate, stability and ligation product formation. Additionally, the premature removal of the auxiliaries from the auxiliary peptides shows that cleavage under mild conditions is feasible and hence, it should be possible to cleave the auxiliaries from the ligation product. For this reason the next stage, finding removal conditions for the auxiliaries from ligation product, was investigated with auxiliary **3.021** and **3.028**.

### 3.1.4 Removal of Auxiliary

The ligation studied with the model peptides have shown that the auxiliaries do indeed seem to be labile to basic conditions as hinted by the auxiliaries' premature removal from the auxiliary peptides. With the observation that the amide bond increases the stability between the auxiliary and the peptide a higher pH value and possibly higher temperature might be necessary to achieve efficient cleavage of the auxiliary.

First tests were performed using either pure ligation product **3.054** (or a mixture of **3.054** and the N- and S-acetylated ligation product **3.054a)**. The peptides were dissolved in an aqueous buffers of 50 mM glycine to a concentration of 2 mM and the pH was adjusted to 9.0, 10.0 and 10.5 with 0.2 M NaOH solution. The mixtures were incubated at 50 °C and 60 °C for 5 hr. A single sample at pH 10.0 was also tested at 80 °C. As can be seen in figure 3.12, no desired product was detected. Observed masses corresponded to disulfide formation and thioester hydrolysis stemming from hydrolysis of peptide **3.054a** (the S-acylataled ligation product). At 80 °C reaction temperature, peaks with unidentified masses were detected- most possibly fragmentation of the peptide induced by high temperature and pH values. The disulfide proved to be a stable product and did not react further. The stability of the disulfide towards auxiliary cleavage points towards the involvement of the thiyl radicaly in the radical pathway (scheme 2.1).

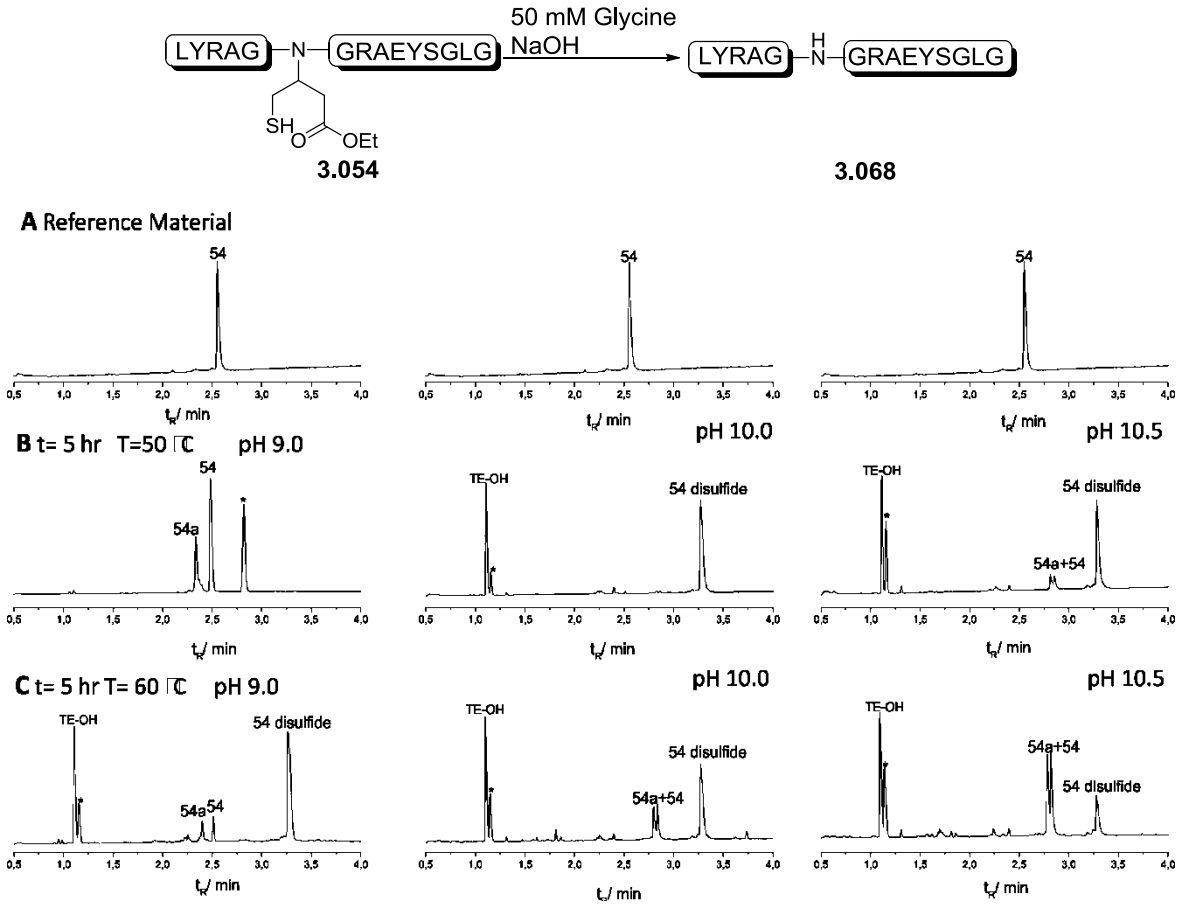



**Figure 3.12**: UPLC analysis of reaction of peptide **3.054** and **3.054a** under treatment with auxiliary removal buffer at different temperature and pH values after 5 hr. **A)** Starting material reference; **B)** 50 °C pH 9.0 (left), pH 10.0 (middle), pH 10.5 (right); **C)** 60 °C pH 9.0 (left), pH 10.0 (middle), pH 10.5 (right); **TE-OH**: LYRAG-OH, peaks labeled with * correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

With the observation of a stable disulfide species, experiments were repeated adding 20 mM TCEP to the removal buffer. In addition, three amine bases were tested: piperidine, morpholine and ethyldimethylamine (EDMA). Aqueous solution containing 160 mM amine and 40 mM TCEP with pH value of 8.5-9.0 depending on the amine were prepared. Peptide **3.054** was dissolved to a concentration of 2 mM and incubated at 80 °C. As seen in figure 3.13 the treatment with NaOH and EDMA failed to yield the desired native peptide. Morpholine and piperidine both gave positive results yielding desired native peptide **3.068**. Importantly, no aminolysis of the ester was observed.

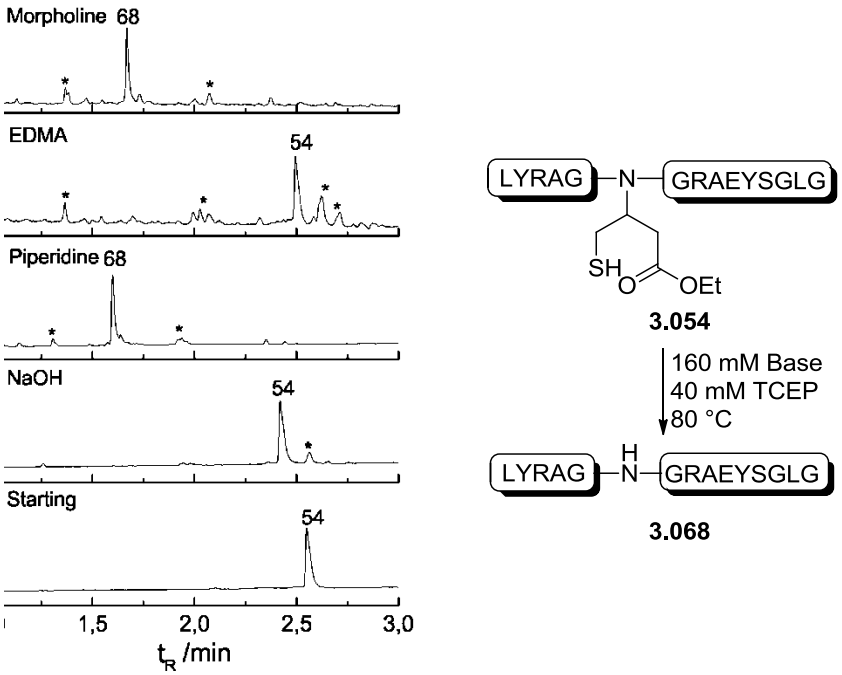
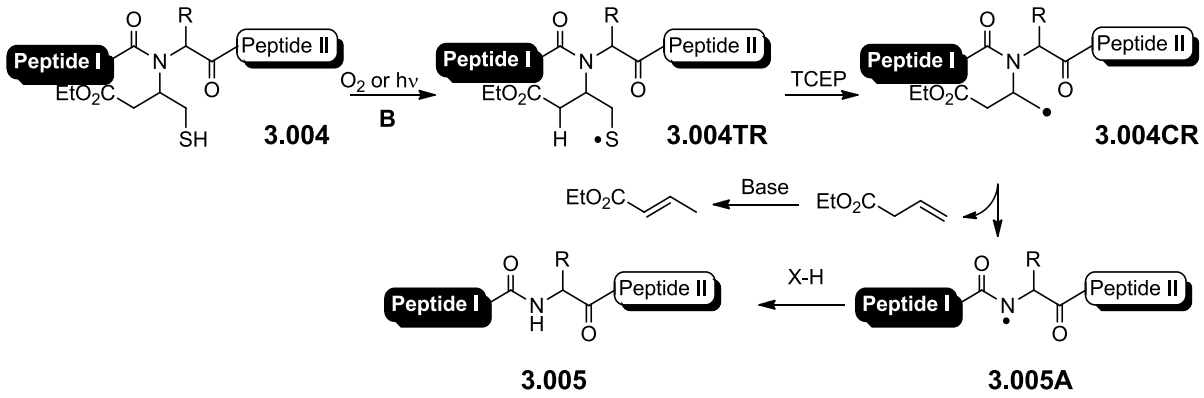



**Figure 3.13**: UPLC analysis of the reaction of peptide **3.054** under different auxiliary removal buffers to yield native peptide **3.068** after 5 hr; Peak * correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

The removal of the auxiliary could be achieved via the previously suggested ionic mechanism, making removal from the amide bond possible by increased temperature and pH. However, the observation that it is the less basic morpholine and piperidine bases which showed the ability to cleave the auxiliary, raised a suspicion that the removal of the auxiliary from the ligation product does not proceed through an ionic mechanism, as was probably the case for the premature removal of the auxiliary. Several observations support a different pathway is for the removal of the auxiliary from the ligation product.

The conditions for the removal, i.e. secondary amine base and necessity of TCEP as additive suggest a non-ionic pathway. Auxiliary cleavage experiments under same conditions but with addition of thiophenol, a potent radical scavenger, failed to promote auxiliary cleavage. These observations add up to results gained with the 2-mercapto-2-phenyl auxiliary[95] and strongly suggest auxiliary cleavage via a radical mechanism.

The possible mechanism is shown in scheme 3.13. Light or oxygen could trigger formation of a thiyl radical **3.004TR**. TCEP induced desulfurization leads to alkyl radical **3,004AR**, which fragments to give the amide radical ligation product **3.005AmR** and the cleaved auxiliary as a Michael acceptor. The amide radical could then be scavenged by the poorer amine hydrogen donors such as piperidine or morpholine.




**Scheme 3.13**: Suggested mechanisms for cleavage of auxiliary from ligation product via a radical mechanism

With positive preliminary results, optimization of the conditions required for the removal was evaluated. Two aspects were tested: the amine component and the temperature. To do so peptide **3.054** was dissolved in a buffer containing 160 mM of either piperidine or morpholine and 40 mM TCEP and incubated at 50 °C or 65 °C (figure 3.14). Analysis of the crude reaction mixture showed removal took place in all cases. Piperidine, being more basic, showed around 70% conversion to product at 50°C and complete conversion at 65 °C after 5 hr. Morpholine exhibited a moderate reaction rate with roughly 50% conversion at 50 °C and 90% at 65 °C after 5 hr. However, morpholine seemed to provide a chemoselectivity as judged from the higher purity of crude materials.

As morpholine proved a milder reagent, forming less side products the piperidine was discarded and temperature optimization experiments were performed. The peptides were incubated at 40, 50 and 60 °C and checked at 2, 4, 6 and 24 hr. As seen in figure 3.15, the auxiliary is removable at all temperatures, however, the removal rate is drastically affected. At 40 °C product formation was first observed after 6 hr and complete conversion achieved after 24 hr. In comparison, after 6 hr at 60 °C reaction was >95% complete. Shorter exposure to the buffer and elevated temperature resulted in a higher crude purity.

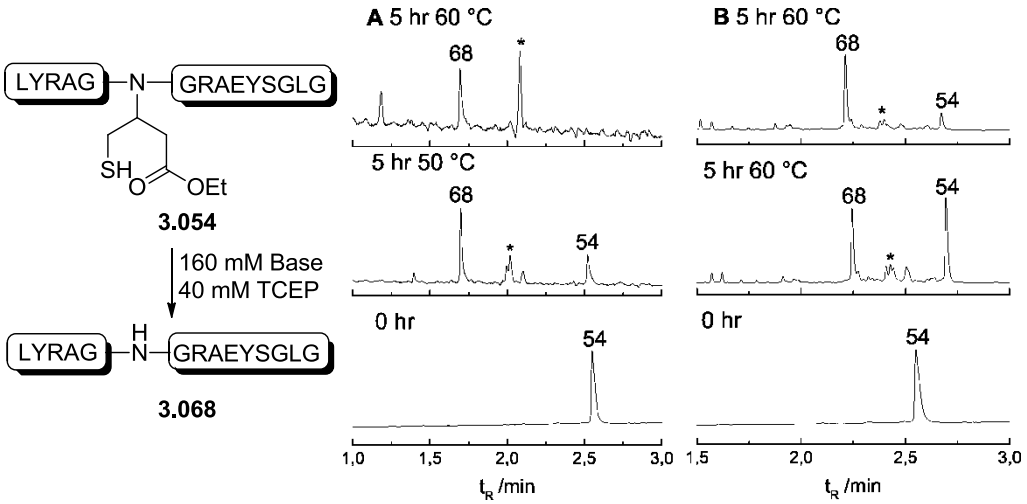


**Figure 3.14**: UPLC analysis of peptide **3.054** under treatment with auxiliary removal buffers containing piperidine (**A**) or morpholine (**B**) at different temperatures to yield native peptide **3.068** after 5 hr; Peak labeled with * correspond to unassigned masses; 03-60% B in 6 min, 210 nm.

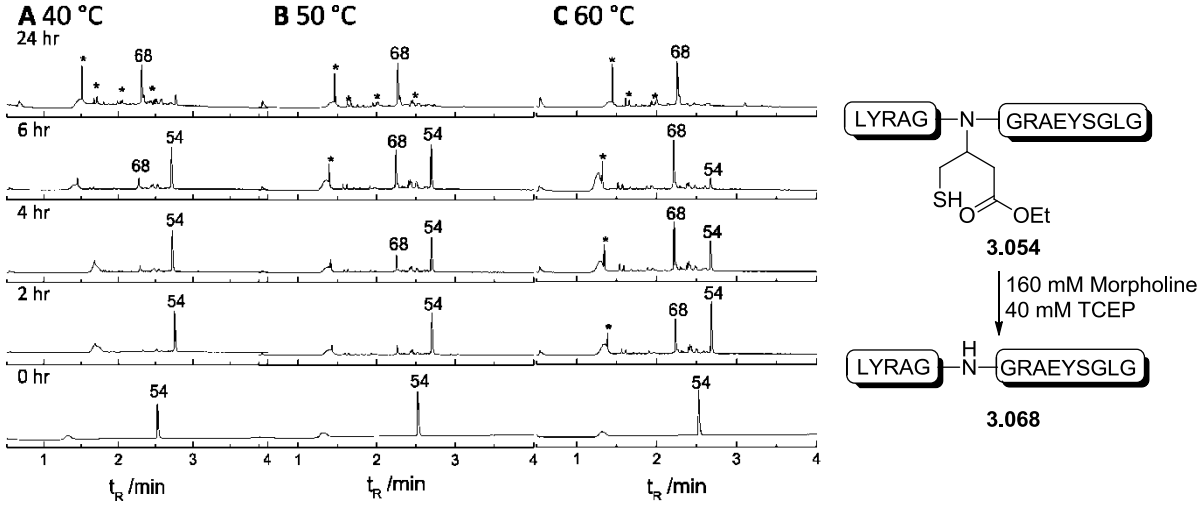



**Figure 3.15**: UPLC analysis showing the advancement of the removal of auxiliary from ligation product **3.054** to give native peptide **3.068** at 40 °C (**A**), 50 °C (**B**) and 60 °C (**C**). Peak labeled with * corresponds to unassigned masses; 03-60% B in 6 min, 210 nm.

Finding removal conditions that promoted mild and selective removal of the auxiliary, the reaction was repeated on ligation products **3.054**, **3.055** ($CO_2Et$), **3.060** and **3.061** (CONHPh) on a preparative scale. The peptides were dissolved in a buffer of morpholine and TCEP and incubated at 60 °C. After 6 hr complete conversion was observed from ligation products **3.054** and **3.055** and after semi-preparative HPLC purification peptides **3.068** and **3.069** were isolated with 43% and 45% yield respectively (figure 3.16).

Compounds **3.060** and **3.061** required longer treatment of between 7-9 hr (figure 3.17). The removal of N-phenylamide auxiliary **3.28** also resulted in lower crude purity with numerous unidentified peaks forming. The unmodified peptides **3.068** and **3.069** were isolated with 35% and 30% yield.

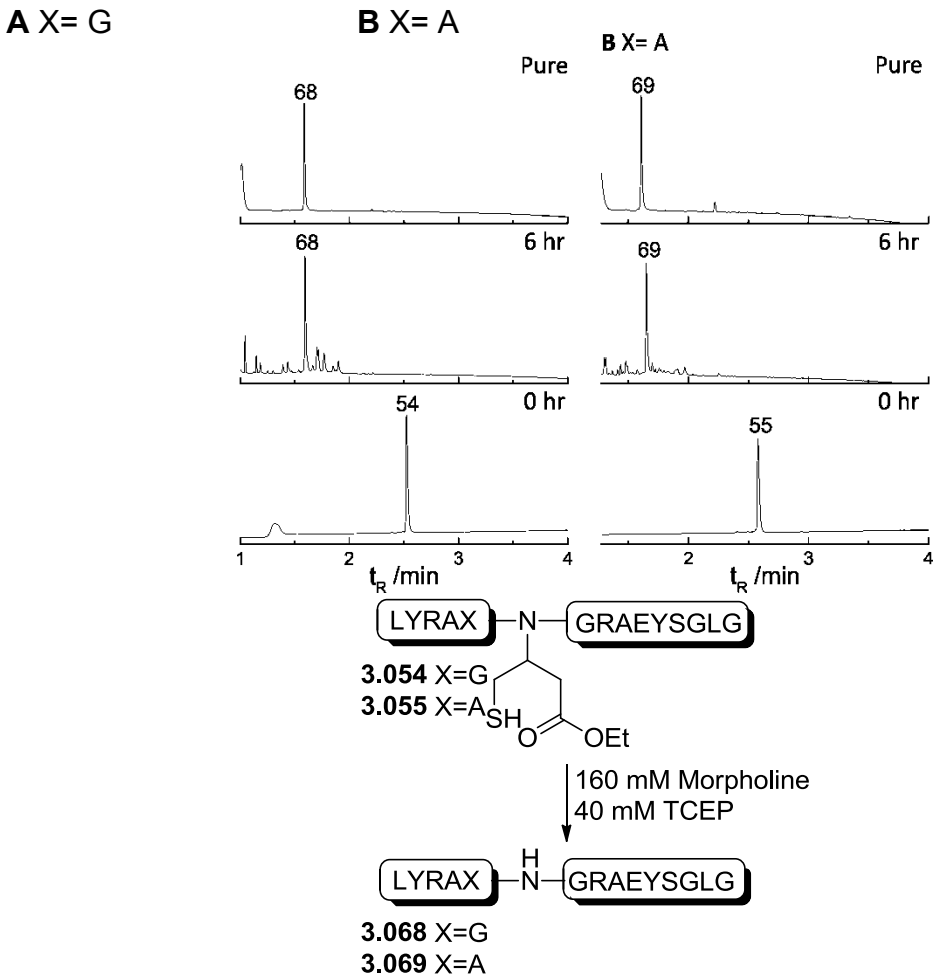



**Figure 3.16**: UPLC analysis showing the removal of auxiliary from ligation product **3.054** to give native peptide **3.068** (**A**) and from ligation product **3.055** to give native peptide **3.069** (**B**) at 0 and 6 hr; Gradient: 03-60% B in 4 min, 210 nm.

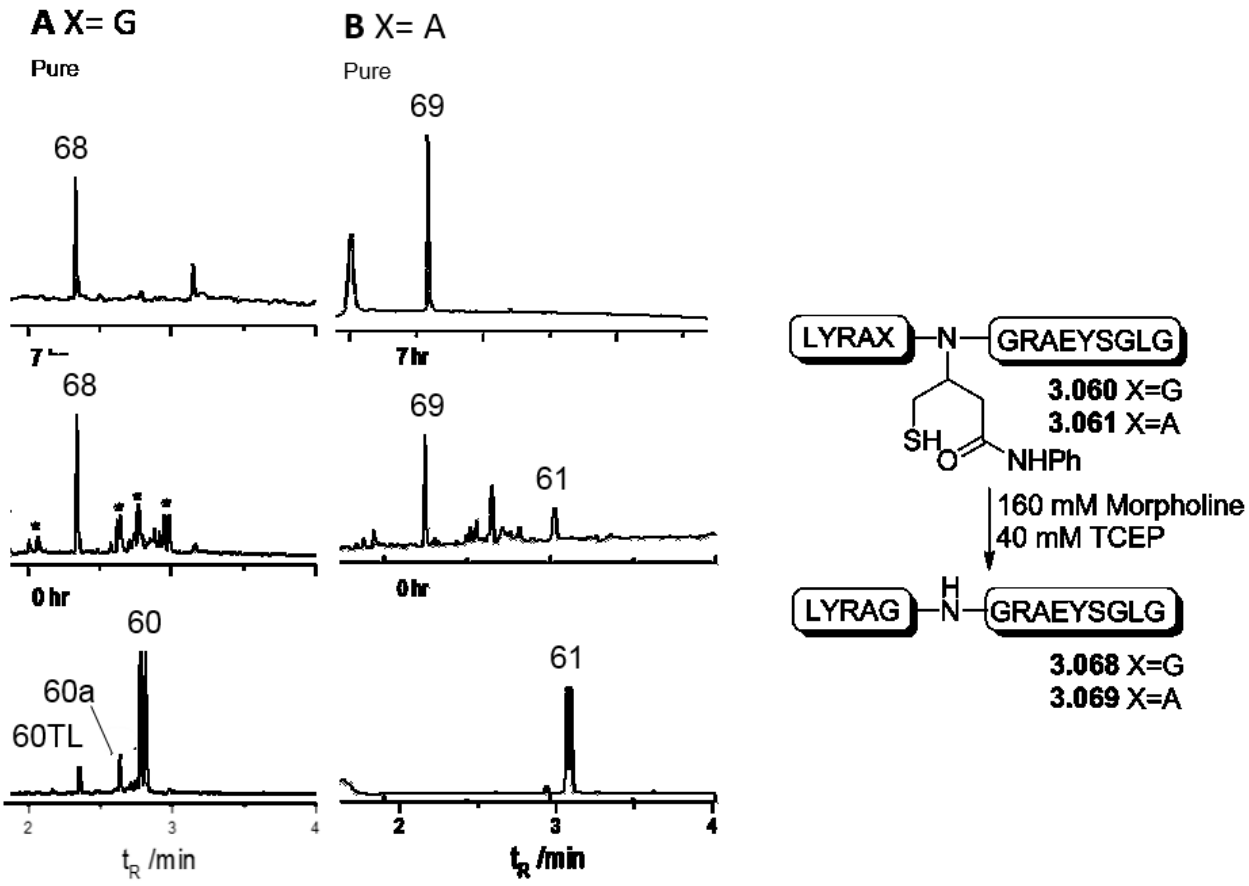


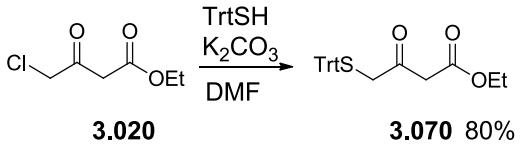

**Figure 3.17**: UPLC analysis showing the removal of auxiliary from ligation product **3.060** to give native peptide **3.068** (**A**) and from ligation product **3.061** to give native peptide **3.069** (**B**) at 0 and 6 hr; Gradient: 03-60% B in 4 min, 210 nm.

The experiments have shown both ethyl ester auxiliary **3.021** and N-phenylamide auxiliary **3.028** can be cleaved from ligation product to yield the unmodified peptide. However, auxiliary **3.021** gave better yields and produced fewer side products during removal. In addition it has a simpler and straightforward synthesis making it the favorable option to use between the two auxiliaries. For these advantages over auxiliary **3.028**, it was decided to continue to the final step of the study, the synthesis of a full-length biologically relevant target, using auxiliary **3.021** only.

## 3.1.5 Changing the Thiol Protecting Group for Improved Protecting Groups Chemistry

The tert-butyl protecting group used during the model studies has several disadvantages. Besides having a foul odor it is stable to treatment with TFA and thus, requires an extra deprotection step by treatment with the highly corrosive TFMSA or mercury acetate which is highly toxic. The added treatment also reduces yield of isolated products and takes longer time to have the final compound in hand. Before pursuing the synthesis of a full length protein an optimized synthesis route was sought after. Two protecting groups stood out as possible alternatives; the triphenylmethyl group (Trt), removed with 90% TFA treatment and alkyldisulfides which are removed under reductive conditions

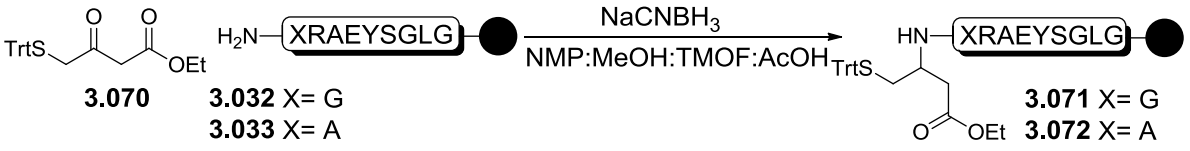
used in ligation. From the two, the triphenylmethyl group was favored for its straightforward synthesis route, as the precursor molecule triphenylmethylmercaptan is commercially available.




**3.020** **3.070** 80%

**Scheme 3.14**: Synthesis of ethyl 3-oxo-4-(tritylthio)butanoate auxiliary precursor **3.070**

The trityl protected auxiliary precursor **3.070** was synthesized in a straightforward manner similar to the synthesis of the tert-Butyl protected auxiliary precursor **3.021** and was isolated with 80% yield (scheme 3.13). Coupling to resin bound peptide **3.32** and **3.33** using the same conditions used for the previous auxiliary precursors (i.e. 10 eq. of auxiliary in NMP, methanol and trimethylorthoformate (TMOF) with 5% v/v AcOH) failed to yield the desired product, most probably due to the increased bulkiness of the Trt group. To push the reaction 50 eq. of the auxiliary had to be added to the resin for the glycine peptide **3.32** and 75-100 eq. for the alanine peptide **3.33**.




**3.070** **3.032** X= G
**3.033** X= A

**3.071** X= G
**3.072** X= A

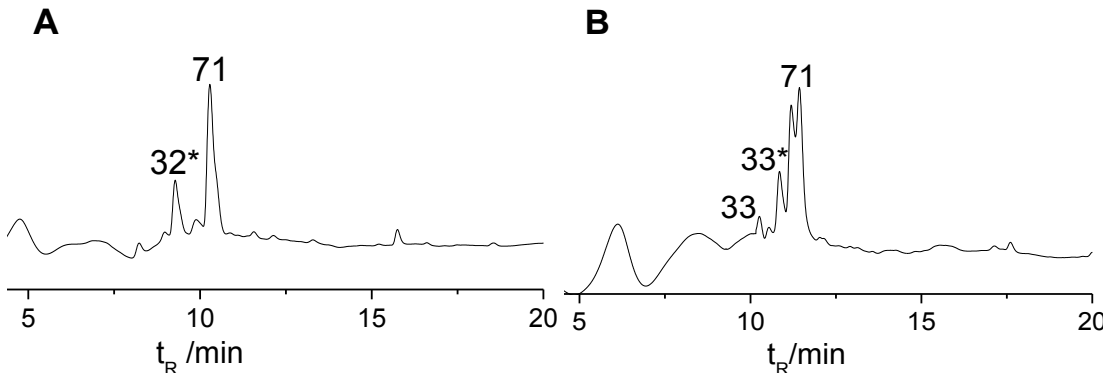



**Figure 3.18**: HPLC analysis showing the solid support reductive amination reaction of compound **3.070** after 48 hr. **A)** Reaction with peptide **3.032**; **B)** Reaction with peptide **3.033**; **3.032***: N-acetylation of GRAEYSGLG **3.032**, **3.033***: N-acetylation of peptide **3.033**; Double peak **3.072** corresponds to the two diastereomers; 03-60% B in 20 min.

## 3.1.6 Total Chemical Synthesis of DCD-1L

### 3.1.6.1 The Dermicidin System

Dermicidin is a 101 residues long protein encoded by the DCD gene. While normally restrictedly expressed in eccrine sweat glands of the skin [103] and in certain parts of the brain [104], overexpression of dermicidin is reported in multiple human tumor types such as cutaneous tumors, melanoma, pancreatic, breast, prostate, and hepatocellular carcinomas [104-111]. The 11 kDa full length dermicidin protein and proteolytic peptides derived from it have been proposed to have diverse biological functions including growth and survival factor in breast cancer [104] and in neural cells [103, 112-115], displaying antibacterial activity [103, 114, 115], and inducing cancer associated cachexia [116, 117]. A recent study also demonstrated that dermicidin may function as a proteolytic enzyme which can promote tumor cell invasion [118].

MRFMTLLFLTALAGALVCAYDPEAASAPGSGNPCHEASAAQKENAGEDPGLARQAPKPRKQR**SSLLEKGLDGAKK AVGGLGKLGKDAVEDLESVGKGAVHDVKDVLDSVL**

**Figure 3.19**: Survival evasion peptide domain and antimicrobial peptide domain sequence of Dermicidin.

Dermicidin belongs to a family of antimicrobial peptides (AMPs), also known as host-defense peptides. This family of peptides plays an important role in the innate immune defense of epithelial barriers. AMPs show a broad activity against a wide range of photogenes including gram positive and negative bacteria, fungi and envelope viruses[119]. It is processed into mature peptides, the most important are DCD-1L, DCD-1 and SSL25. It shares no homology with other known AMPs and is an unstructured random coil at its secreted state. Unlike most AMPs which are cationic, DCD-1L is unique for being anionic with a net charge of -2 at physiological pH.

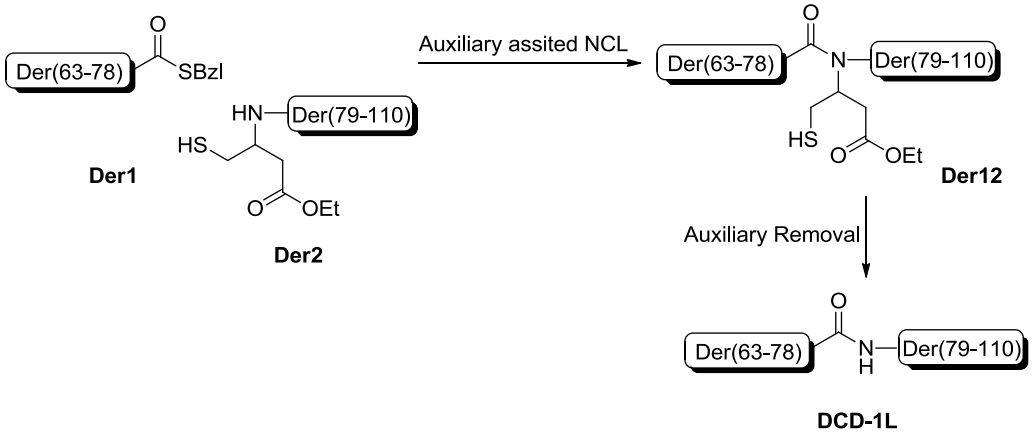
Giving the AMPs important role in the immune response as well as their involvement in several inner cellular signal transduction pathways, a method to produce pure homogenic samples of DCD-1L and other AMPs can help the field of cell biology, immunology, chemical biology or microbiology in studying these systems. As many antimicrobial proteins lack a cysteine residue but possess several Gly-Gly junctions [120], auxiliary-mediated NCL can be a helpful tool for synthesis of large amounts or libraries of AMPs. Herein the presented ethyl 4-mecaptobutyrate auxiliary **3.070**, which is readily synthesized from inexpensive, commercially available starting materials and is easily introduced directly onto the polymer bound peptide, was utilized for the synthesis of the 48 residues long DCD-1L.

### 3.1.6.2 Synthetic Route towards DCD-1L

Scheme 3.15 shows the sequence and synthetic strategy towards the total chemical synthesis of the 48 amino acid long DCD-1L peptide. The peptide contains no cysteine residues but possesses a Gly[78]-Gly[79] junction making it an ideal target to test the auxiliary-assisted native chemical ligation strategy.

The thioester fragment **Der1** was synthesized employing the safety-catch linker strategy[121] (figure 3.20). In this method the peptide was synthesized according to the Fmoc strategy on a sulfonamide linker. After completion of the synthesis the sulfonamide was activated by treatment with trimethylsilyldiazomethane. The peptide was cleaved from support by thiolysis with benzyl mercaptan and protecting groups were removed by treatment with TFA to yield the crude unprotected thioester. The crude was purified on a preparative HPLC to yield **Der1** in 35% yield.

SSLLEKGLDGAKKAV**GG**LGKLGKDAVEDLESVGKGAVHDVKDVLDSVL




**Scheme 3.15**: Sequence and synthetic route towards DCD-1L.

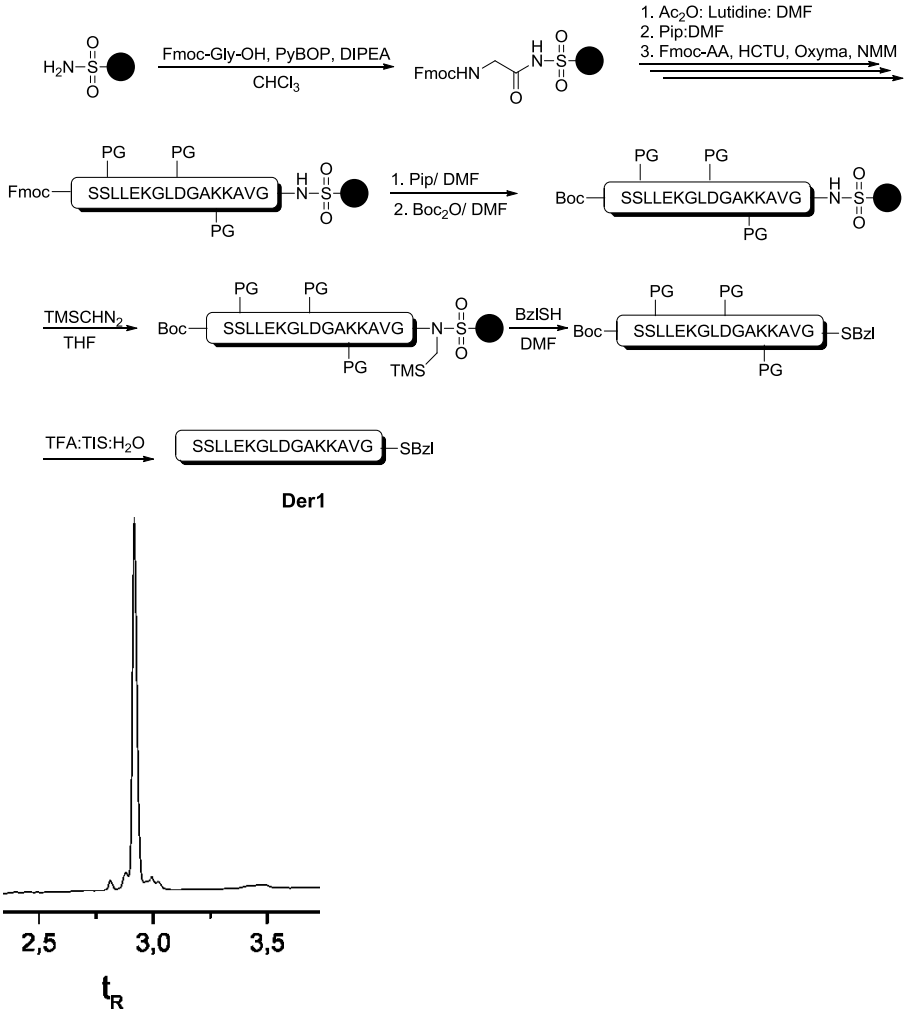



**Figure 3.20**: Synthesis of **Der1**. UPLC-MS analysis of pure **Der1**. 03-60% B in 6 min, 210 nm.

**Der2** was prepared employing Fmoc strategy on a Rink Amide resin. After synthesis was complete auxiliary precursor **3.69** was attached directly onto the resin bound peptide using the previously described reductive amination on solid support method. A significant amount of acetylated peptide was detected, reducing the yields of the final product. Peptide was cleaved and all protecting groups removed using standard TFA cleavage and crude was purified on a preparative HPLC to yield the pure peptide **Der2** in 10% yield (figure 3.21).

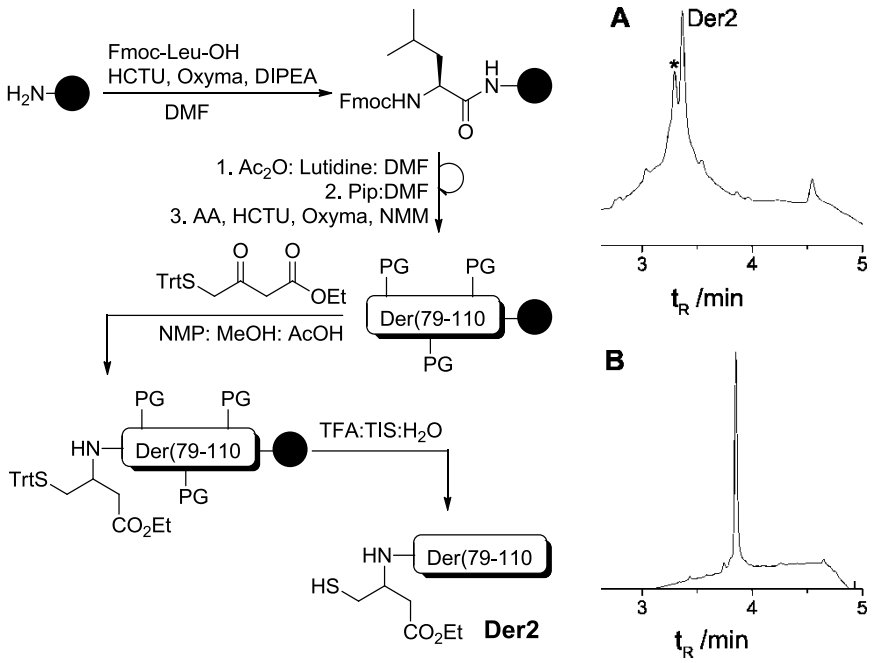



**Figure 3.21**: UPLC analysis of crude (**A**) and pure (**B**) Aux-Der(78-110) **Der2**. Peak * corresponds to acetylation of amine. Gradient crude: 03-90% B in 6 min; Gradient pure: 03-70% B in 6 min, 210 nm.

### 3.1.6.3 Total Chemical Synthesis of DCD-1L by Auxiliary-Assisted Native Chemical Ligation

The ligation between **Der1** and **Der 2** was carried out in a similar fashion to the model study. The peptides were dissolved in a phosphate buffer containing 20 mM TCEP, 2% v/v thiophenol to a concentration of 5 mM and incubated at 25 °C. After 7 hr ligation was complete and the ligation product **Der12** was isolated in 50 % yield. To remove the auxiliary a protocol optimized by Simon Loibl (ongoing thesis work, Humboldt University Berlin) was applied, using higher concentrations of morpholine and TCEP. Pure ligation product **Der12** was dissolved in aqueous buffer containing 400 mM morpholine and 100 mM TCEP to a concentration of 1-2 mM and incubated at 50 °C for 6 hr after which UPLC-MS analysis showed a complete removal of the auxiliary. The native peptide was purified on a semi-preparative HPLC to yield the pure **DCD-1L** peptide with 40 % yield (figure 3.22).

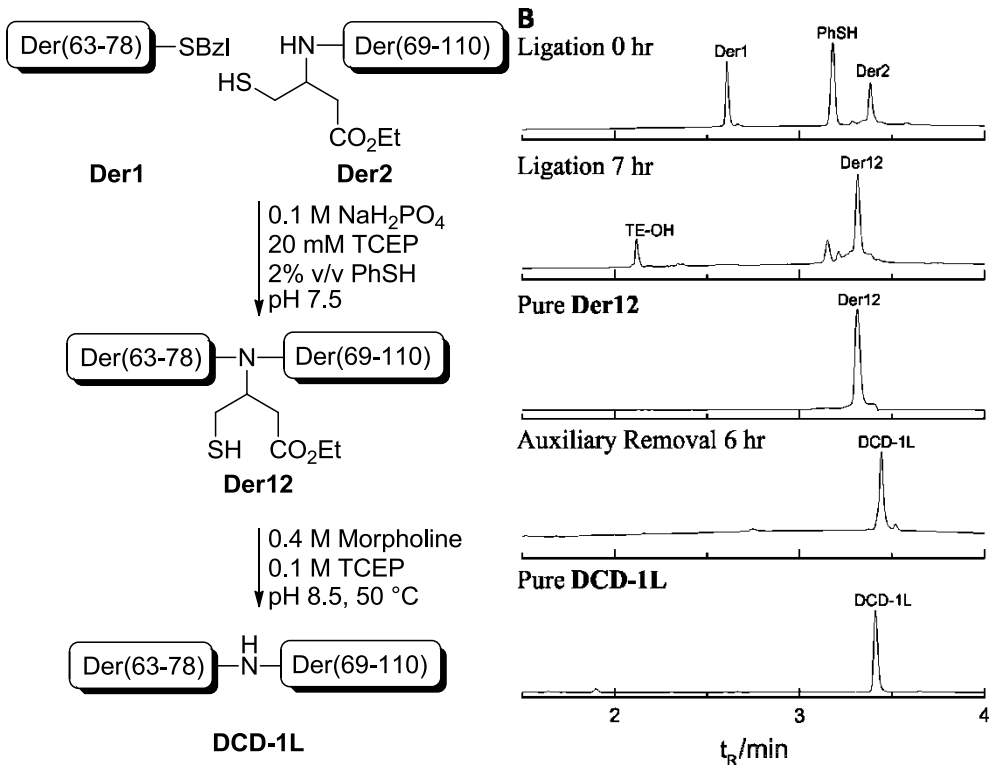



**Figure 3.22**: Synthesis of **DCD-1L**. **A)** Scheme depicting synthetic route; **B)** UPLC analysis of ligation between **Der1** and **Der2** and auxiliary removal to yield native **DCD-1L**.

## 3.1.7 Conclusions

In this work a new class of $N^{\alpha}$-auxiliaries for auxiliary mediated NCL was described. The auxiliaries were envisioned to advance the field in several aspects:

1. Broaden the scope of synthetic targets available through auxiliary assisted NCL by designing a flexible scaffold that would allow ligation in sterically demanding junctions. As previously published auxiliaries failed to promote ligation at sterically demanding junctions the field of auxiliary mediated ligation did not mature to the level of some other methods such as the desulfurization method which enjoy high usability and important application.
2. Develop an auxiliary that could be removed under mild basic conditions, avoiding the acidic conditions used in the available auxiliaries. Acidic conditions induce an N→S acyl shift, reversing the ligation direction and more importantly are incompatible with a variety of post translational modifications, a fast growing field with its impact and importance only beginning to unravel.
3. Inexpensive, readily synthesized auxiliaries and straightforward coupling to a peptide. While not detrimental, such features would increase the applicability and generality of the method, enlarging its penetration and application to the scientific community.

The ligation studies on model peptides have shown that the auxiliaries developed were unable to promote ligation at the Ala-Ala junction. This result shows that despite using a linear, flexible scaffold

the amine was still too hindered as the amine has two branched alkyl substituents, blocking the S→N acyl shift.

A second disadvantage of the auxiliaries is their instability. Under conditions of NCL the auxiliary peptides were unstable and the auxiliary was cleaved to produce native peptides GRAEYSGLG **3.032** or ARAEYSGLG **3.033**. This undesired side reaction varied between the four auxiliaries based on the auxiliary tendency to form the thiolactone species with amide auxiliary **3.017** and N-methyl-N-phenyl auxiliary **3.019** being especially prone to cyclization.

Ethyl ester auxiliary **3.016** and N-phenylamide auxiliary **3.017**, while not unsusceptible to cyclization, are less prone to form the thiolactone making these auxiliaries usable for ligation junctions of the X-Gly family and highly efficient at Gly-Gly junctions. This efficiency was demonstrated by synthesizing the full-length antimicrobial peptide DCD-1L.

Advantages of auxiliary **3.016** include simple synthesis of the auxiliary precursor from commercially available and inexpensive starting materials and effective attachment onto a peptide directly on solid support contributing to the generality of the method by obsoleting the need to synthesize specific building blocks.

The major achievement gained with these auxiliaries is the ability to cleave them from the ligation product under mild basic conditions. After optimization of the conditions needed for the removal, a mild and fast procedure for the cleavage of the auxiliary from the ligation product was successfully developed. As basic conditions were applied for removal of the auxiliaries, they should prove especially attractive when synthesis of proteins with acid sensitive posttranslational modification such as sulfonation or glycosylation are pursued. These advantages establish auxiliary **3.016** as a legitimate and worthy addition to the existing ensemble of auxiliaries.

# 3.2 Synthesis of MUC1 Protein by Auxiliary-Assisted Solid Phase Chemical Ligation

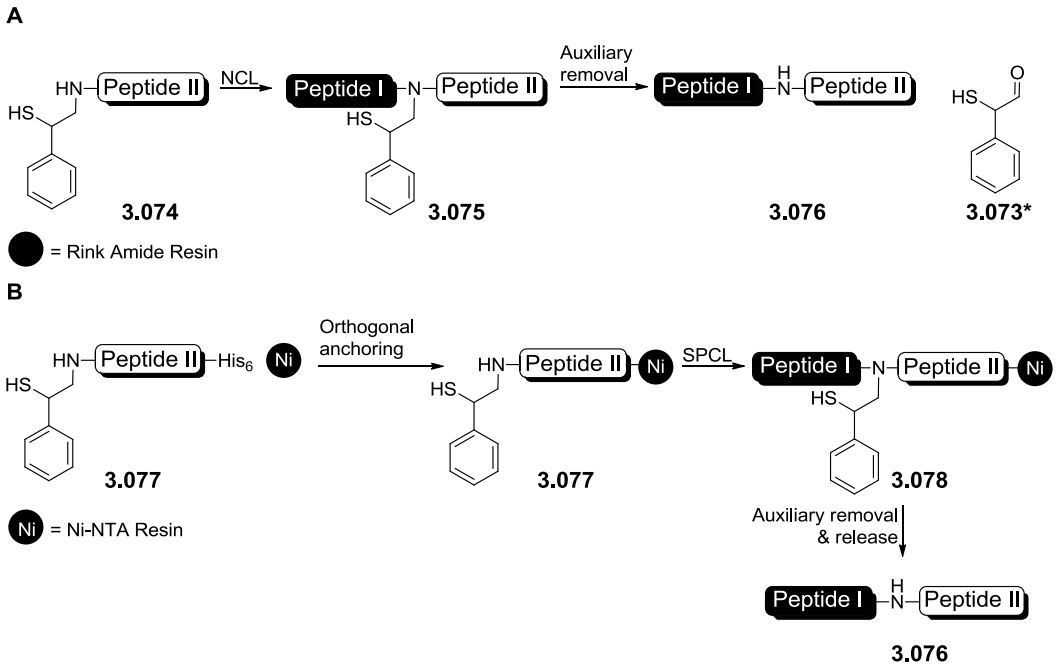
## 3.2.1 Concept and Introduction

Seitz and coworkers recently reported a new class of highly potent $N^{\alpha}$-auxiliary that has shown improved characteristics during ligation studies [95]. The 2-mercapto-2-phenylethyl auxiliary **3.073** (scheme 3.16) can promote ligation at junctions past the Gly or Ala barrier. The precursor **3.073*** (scheme 3.16 A) is readily attached onto a resin bound peptide by reductive amination thanks to its aldehyde moiety, making it a universal tool for preparation of auxiliary modified peptides. Following ligation the auxiliary is selectively cleaved using mild conditions directly from the ligation mixture without intermediate purification thus reducing costly and lengthy purification steps.

NCL continue to lead as the method of choice for synthesis of protein. Done in solution, it enables easy analysis of the reaction at every given time and can be directly purified as soon as the ligation is complete. Despite these advantages, attempts are constantly being made to apply ligation chemistry to solid phase. The motivation behind these trials is to exert the advantages SPPS gave the field of peptide synthesis and lend them to the field of total chemical synthesis of proteins.

Despite progress in the field of SPCL, it still has not become a default method in the protein chemist's toolbox. The reasons are compatibility issues of NCL and its extensions to the added requirements dictated by the polymer support, lack of a variety of suited polymer supports, and orthogonality problems between the reagents needed for the different stages (such as: anchoring and releasing of the peptides onto and off the support or protecting groups removal).

The 2-mercapto-2-phenylethyl auxiliary **3.073** with its mild conditions for removal and ability to ligate broader range of junctions has the potential to make SPCL common practice. The use of commercially available Ni-NTA resin (Nickel Nitriloacetic acid), simplifies the setting of the reaction as the peptides are readily attached to the support (scheme 3.16B).




**Scheme 3.16**: Auxiliary assisted chemical ligation using auxiliary **3.73** in solution (**A**) and on solid support (**B**).

## 3.2.2 The MUC1 Protein

MUC1 is part of the mucin family proteins. These family of proteins is characterized by the presence of a mucin domain rich in proline, serine and threonine residues (PTS domain). The MUC1 is expressed in epithelial cells in mammals and serves many functions. It participates in numerous metabolic processes in the cell [122], provides protection against pathogens [123] as well as act as lubricant to protect the ephitelia [124]. It is a transmembrane protein consisting of three main domains (figure 3.23 A). The C-terminal consists of two domains: a cytoplasmic tail and a transmembrane part. The long extracellular PTS domain, also called variable tandem repeat (VNTR) makes up the N-terminal of the protein. It is constructed of multiple repeats (ranging between 20-120 repeats) of identical 20 residues long peptide (PDTRPAPGSTAPPAHGVTSA). It is highly O-glycosylated at its numerous serine and threonine residues. The massive oligosaccharide construct is responsible for the protection of the cell by providing lubrication and physical barrier, preventing pathogens from reaching the cell surface.

The MUC1 is overexpressed in most human epithelial cancers. The tumor-associated MUC1 (TA-MUC1) differs from the normally expressed MUC1 (figure 3.23 B and C). Normal MUC1 is heavily glycosylated and localized at the apical surface of the cell. As cancer cells lack polarization and differentiation between different areas in the cell, TA-MUC1 is expressed all around the cell leading to over-expression. The TA-MUC1 is also hypo-glycosylated. These structural and spatial changes have been associated with increased endocytosis, and oncogenic signaling. The processes in which this happens are not completely clear. One theory is that the hypo-glycosylation unmasks the

peptide core thus allowing proteases to cleave the peptide extracellular domain which in turn causes conformational changes in the cytoplasmic tail, activating downstream cell signaling.

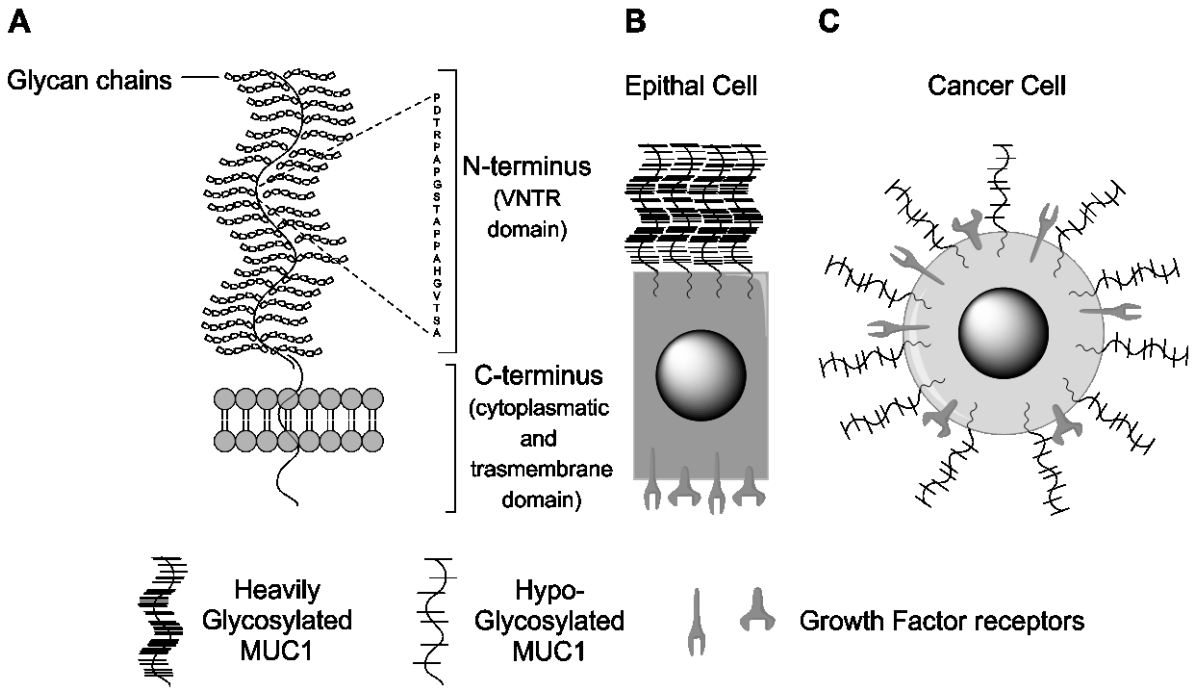



**Figure 3.23**: **A)** MUC1 structure showing the different domains and their location; **B)** MUC1 expression and distribution in healthy epithal cells; **C)** MUC1 expression and distribution in cancer epithal cells.

The numerous roles MUC1 plays in cells and especially in cancerous cells make it imperative to understand and study its multifaceted mechanisms of action. The MUC1 is highly polymorphic in regards to the number of tandem repeats, the tandem repeat sequence variation as well as the glycosylation profile. All vary greatly between different tissues and cells. Obtaining homogenous samples of the peptide can be challenging. Chemical synthesis is a valuable tool for such complicated systems as it can produce homogenous samples of the protein with any desired modification. The MUC1 VNTR domain, with its multiple repeats, makes SPCL especially adequate for its synthesis, as the same "monomer" peptides can be sequentially added multiple times in a continuous, semi-automatic process.

## 3.2.3 Choice of Polymer Support

Polymer support plays a key role in the successful utilization of SPCL. It must swell in aqueous mediums, allow easy attachment and release of the peptide fragments in an orthogonal fashion, and be stable to the reagents used in the ligation chemistry. Several polymeric materials have stood out as being compatible with these requirement. Most notably the PEGA resin has been successfully used for the synthesis of full length protein through the sequential SPCL of up to five peptide segments[94]. One of the limitations of the PEGA resin is to engineer the selective anchoring

of the unprotected peptide. Current methods require derivatization of the polymer and peptide to afford orthogonal binding. In order to avoid extra manipulation steps on peptides or resins to achieve orthogonal anchoring and release from the support, the commercially available Ni-NTA affinity resin was chosen[125]. It consists of beaded agarose derivatized with a nitrilotriacetic acid chelating moiety loaded with nickel ions (figure 3.24). The resin exhibits excellent swelling properties in aqueous medium. Adding a sequence of six histidine residues at the C- or N-termini of the peptide allows efficient and swift anchoring to the support. Detachment of the peptide from the support is also performed under mild conditions by suspending the resin in an aqueous buffer with high concentration of imidazole which competes with the His-Tag for binding the Ni(II)

The Ni-NTA resin has long been used in the chemical and molecular biology fields for affinity chromatography of expressed proteins modified with a $His_6$-Tag, taking advantage of the strong imidazole-nickel complex to selectively extract the desired protein from a cell lysate[126]. Kent et al. have used this resin for the non-chromatographic purification of ligation product[127].

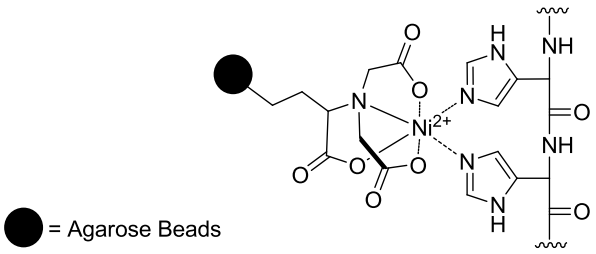
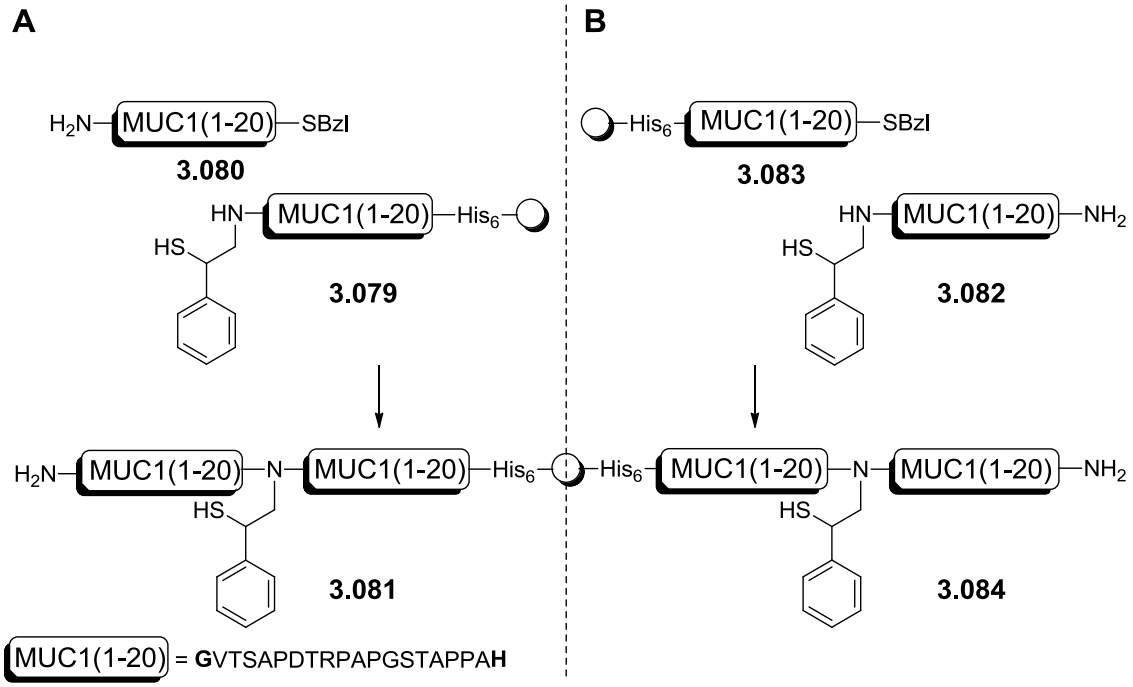



**Figure 3.24**: Ni-NTA resin structure and coordination to histidine residues.

## 3.2.4 Synthesis of Peptide Fragments for the C→N and N→C directed SPCL

In NCL, which deals with the conjugation of two molecules, a concept of directional synthesis is irrelevant. Only if three or more peptide segments are needed to construct the full-length protein does direction come into play. SPCL, like SPPS or however, is always directional and performed either in C→N or N→C direction (scheme 3.17) depending on the peptide terminus that is attached to the solid support. To test the two options four peptides were synthesized: two auxiliary peptides and two thioester peptides with and without a His-Tag.




**Scheme 3.17**: Auxiliary-assisted SPCL of MUC1in C→N direction (**A**) and in N→C direction (**B**).

Aux-MUC1-$His_6$ **3.079** was synthesized following Fmoc procedure. The auxiliary **3.073** was introduced by reductive amination. After 4 hr UPLC-MS analysis showed complete conversion to the desired auxiliary modified peptide Aux-MUC1(1-20)-$His_6$ **3.079**. The peptide was cleaved from support by TFA treatment and purified by preparative HPLC to give the desired product in 40% isolated yield (figure 3.25).

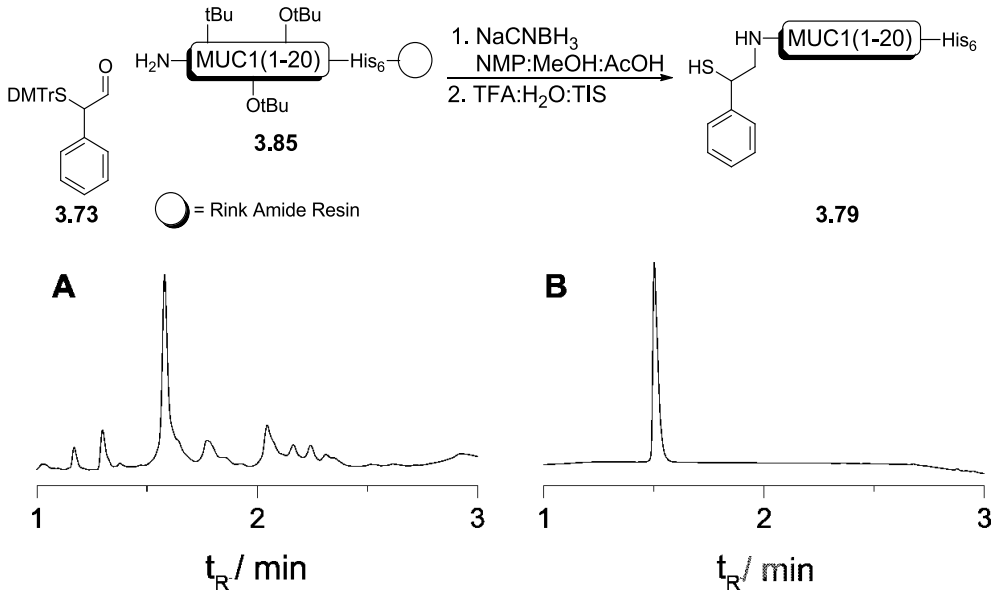



**Figure 3.25**: UPLC analysis of crude (**A**) and pure (**B**) Aux-MUC1(1-20)-$His_6$ **3.079**; 03-60% B in 6 min, 210 nm.

As discussed previously, one of the goals of the project was to perform sequential SPCL. This requires a middle fragment containing the thiol and thioester functionalities. With this outlook in mind, peptide Aux-MUC1(1-20)-$NH_2$ **3.082** was instead synthesized as 3,4-diamino benzoic acid

(Dbz) derivative Aux-MUC1(1-20)-Dbz **3.086**. The Dbz moiety functions as a latent thioester which can be activated post-ligation to function as thioester to allow elongation of the peptide chain (see section 3.2.8).

Aux-MUC1-Dbz **3.086** was synthesized as described in literature for synthesis of Dbz peptides [128] (figure 3.26). Shortly, Fmoc-Gly-OH and Fmoc-Dbz-OH were manually coupled on a rink amide resin. The free amine group of the Dbz was protected as allyloxycarbamate (Alloc) and Fmoc was removed from the second amine. As the protection of one of the amines on the ring deactivates the neighboring amine, Fmoc-His(Trt)-OH was double coupled using HATU to ensure complete acylation of the amine. The rest of the sequence was synthesized by automated SPPS and the auxiliary **3.073** was introduced as described previously. The Alloc protecting group was removed by treatment with tetrakis(triphenylphosphine)palladium(0) and peptide was cleaved from support and protecting-groups removed with TFA:TIS:$H_2O$ cocktail. After preparative HPLC purification peptide was isolated in 30% yield.

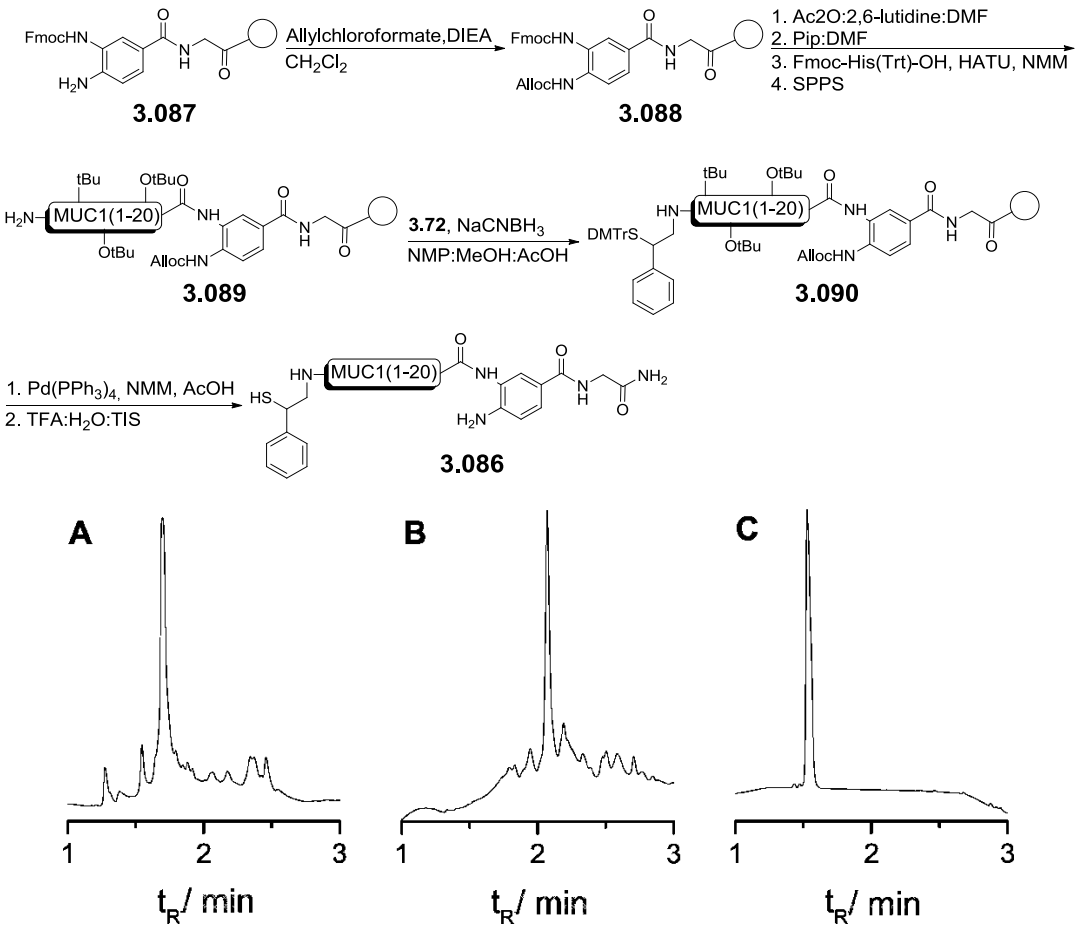



**Figure 3.26**: UPLC analysis of synthesis of Aux-MUC1(1-20)-Dbz-Gly **3.086**. **A**) Crude H-MUC1(1-20)-Dbz(Alloc)-Gly **3.089**; **B**) Crude Aux-MUC1(1-20)-Dbz(Alloc)-Gly peptide **3.090**; **C)** Pure Aux-MUC1(1-20)-Dbz-Gly **3.086**; 03-60% B in 6 min, 210 nm.

The thioesters peptides $His_6$-MUC1(1-20)-SBzl **3.093** and H-MUC1(1-20)-SBzl **3.094** were synthesized using the safety-catch linker approach and isolated with 35% and 45% yield respectively (figure 3.27).

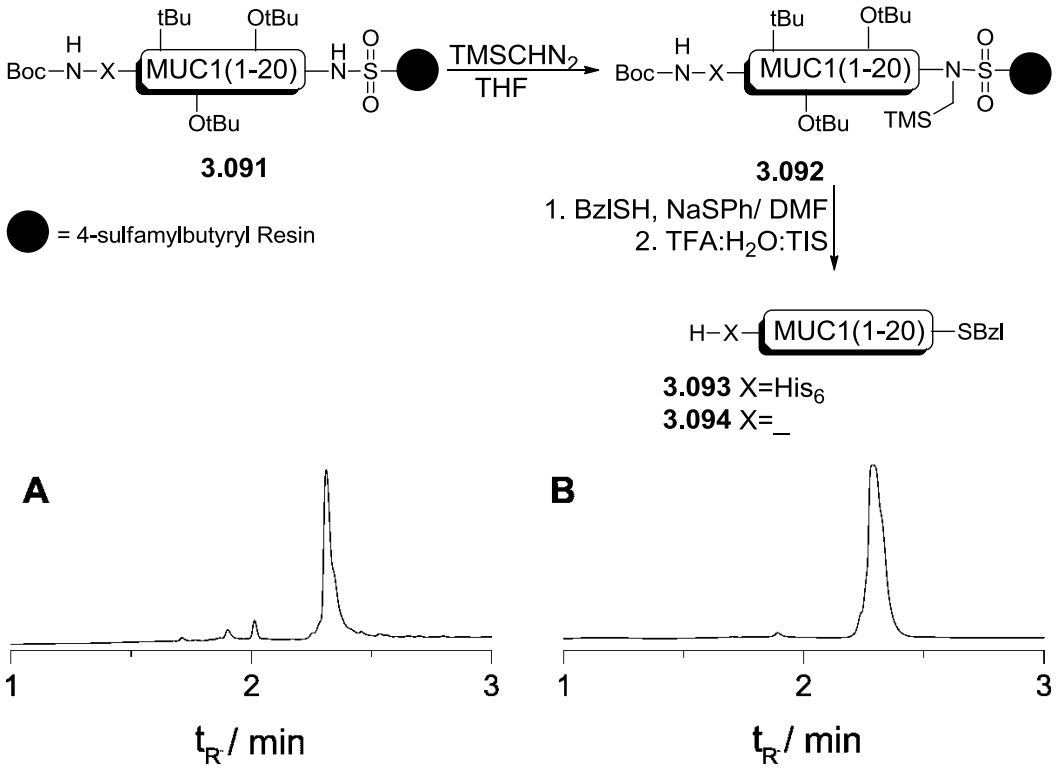



**Figure 3.27**: UPLC analysis **A)** Pure $His_6$-MUC1(1-20)-SBzl **3.093**; **B)** Pure H-MUC1(1-40)-SBzl **3.094**; 03-60% B in 6 min, 210 nm.

## 3.2.5 Attaching the peptides to the polymer support

The His-Tagged peptides **3.079** and **3.093** were attached to a Ni-NTA resin pre-swelled in guanidinium phosphate buffer. After 1 hr full anchoring onto the resin was achieved (figure 3.28). The stability of the system to long exposure of thiols and TCEP was tested over three days (see p. 132) Results have shown the anchoring to the resin was not harmed with no peptide detected in the supernatant.

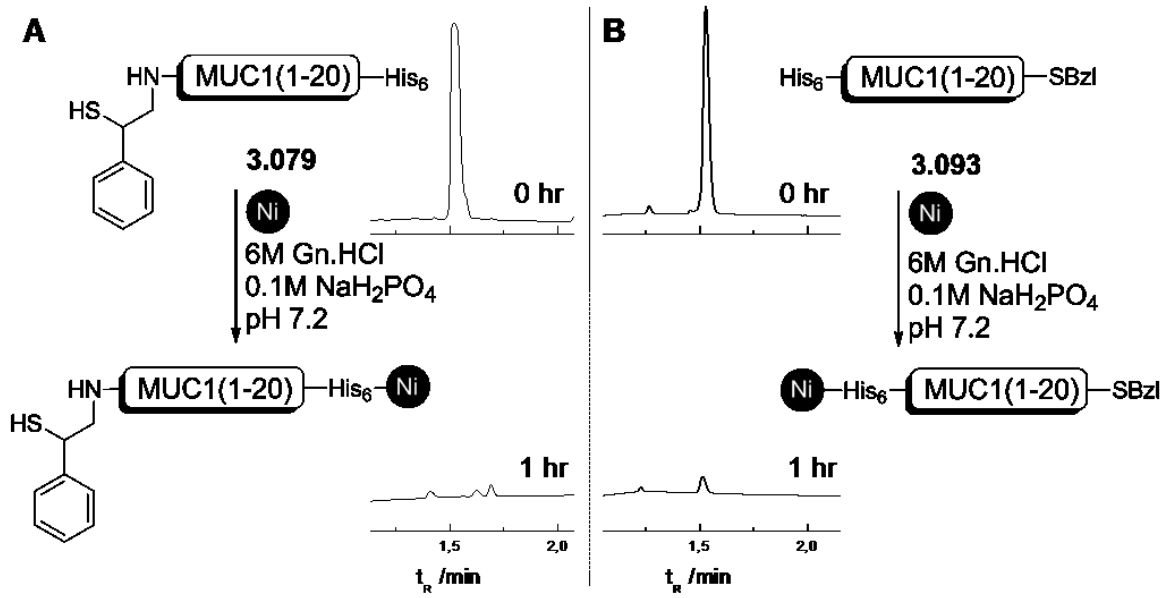



**Figure 3.28**: UPLC analysis of anchoring of peptides to the Nickel-NTA support. **A)** Auxiliary peptide **3.079**; **B)** Thioester peptide **3.093**; 03-60% B in 6 min, 210nm.

## 3.2.6 Auxiliary Assisted SPCL on Ni-NTA Resin

The ligation studies were carried out simultaneously on the two resin bound peptides **3.079** and **3.093** by adding 1.1-1.2 equivalents of thioester **3.094** or auxiliary peptide **3.086** in ligation buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 10 mM TCEP, 10 mM MPAA pH 7.2 and added to the resin. After 16 hr the resin was washed with buffer and water and the attached peptides released by suspending the resin in aqueous solution of 250 mM imidazole and 10 mM TCEP pH 8 for 30 min. The elution solution was collected and analyzed by UPLC-MS. Results demonstrated both directions to be equally effective producing a single peak corresponding to ligation product with crude purity of 85% and 92% for peptide **3.095** and **3.096** respectively (figure 3.29). The two experiments have proven the compatibility of the Ni-NTA resin to SPCL application as well as the auxiliary capability to mediate ligation on solid support effectively and cleanly.

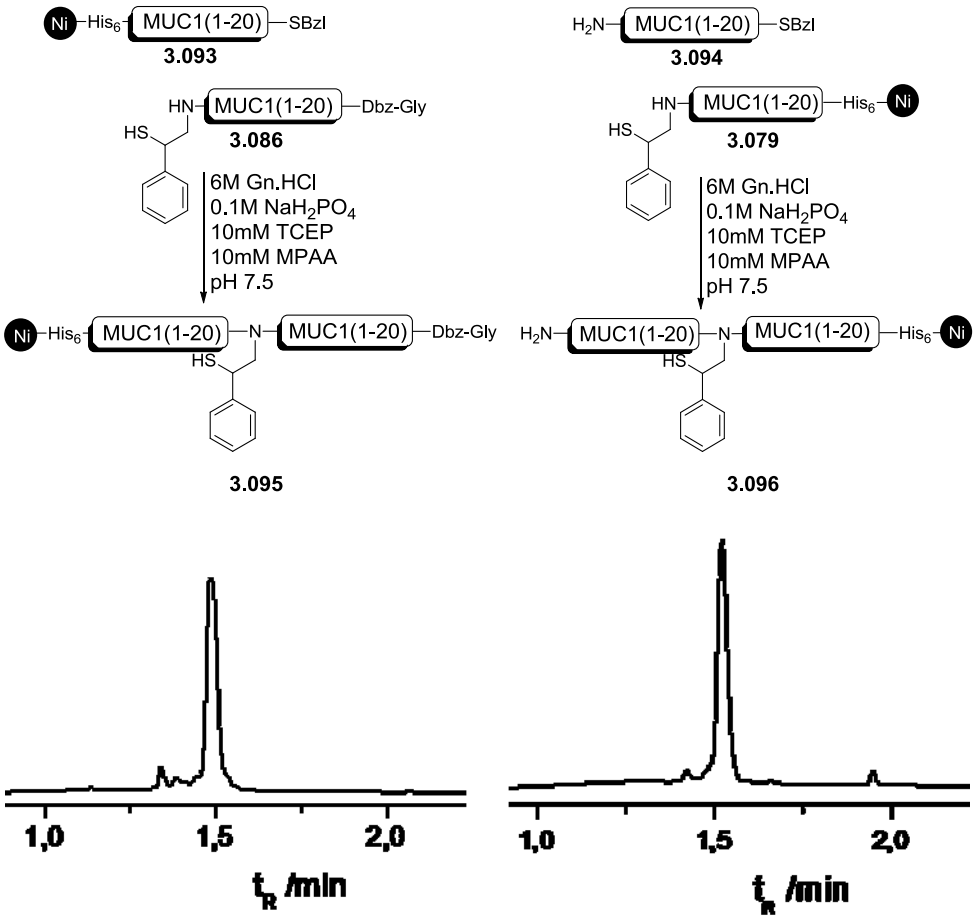



**Figure 3.29**: UPLC analysis of auxiliary assisted SPCL. **A)** N→C directed SPCL to give ligation product **3.095**; **B)** C→N directed SPCL to give ligation product **3.096**; 03-60% B in 6 min, 210 nm.

## 3.2.7 Auxiliary Removal from ligation product

Having proved the auxiliary's ability to mediate ligation on solid support, removal of the auxiliary was tested. The resin bound ligation product $His_6$-MUC1(1-20)-Aux-MUC1(1-20)-Dbz-Gly-$NH_2$ **3.095** was released from support as described previously. To the elution buffer containing the peptide equal volume of auxiliary removal buffer (400 mM morpholine, 100 mM TCEP) was added and reaction was performed at 40 °C for 16 hr.

As can be seen in figure 3.30 auxiliary removal was achieved but not completed. The incomplete removal is probably the result of the lower concentration of morpholine (200 mM) and TCEP (50 mM) combined with the slightly lower pH value due to dilution of the buffer with the imidazole solution.

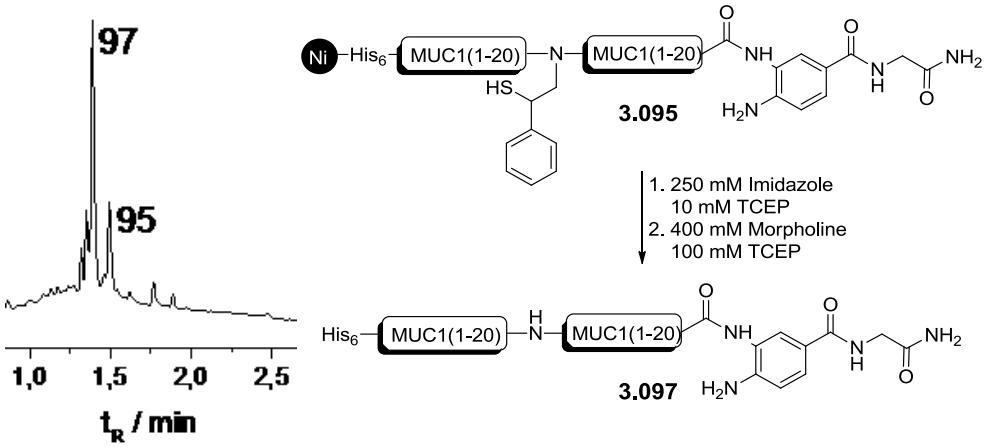
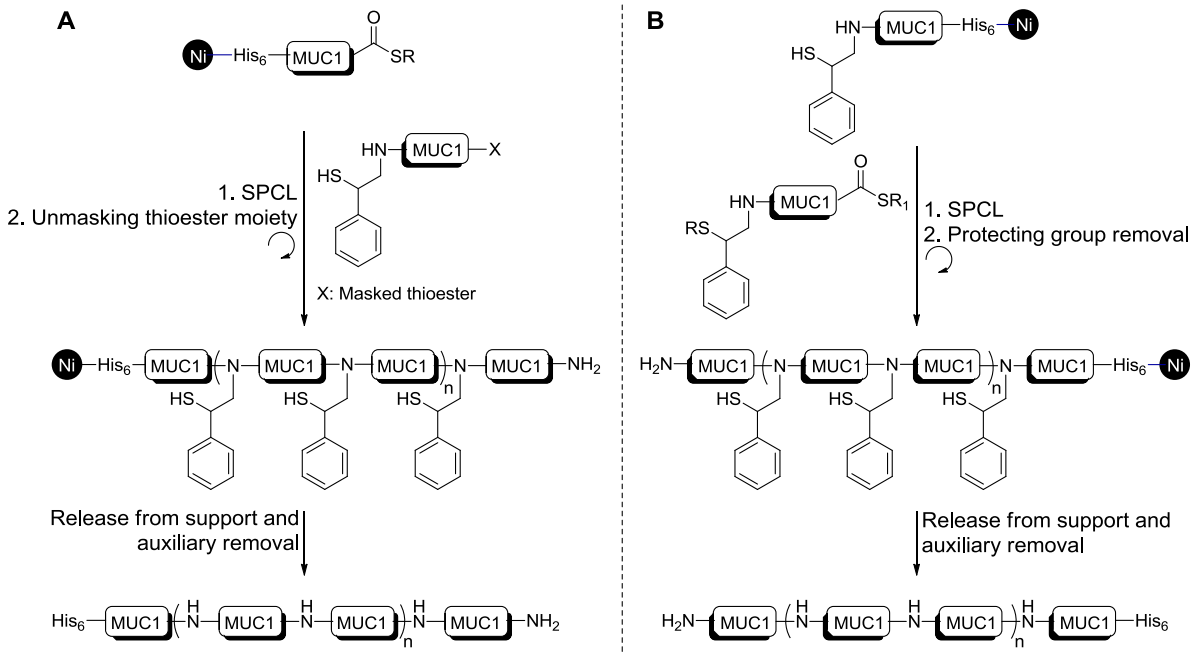



**Figure 3.30**: UPLC analysis of auxiliary removal from ligation product **3.095** after 16 hr at 40 °C; 03-60% B in 6 min, 210 nm.

## 3.2.8 Efforts towards Sequential Auxiliary-Assisted SPCL

With the successful initial studies of auxiliary-assisted SPCL in both C→N and N→C directions, the next goal was performing sequential SPCL using the auxiliary. This requires a synthesis of a middle fragment, having both thioester and auxiliary functionalities. To ensure selectivity one of the functional groups must be temporarily blocked (scheme 3.18).



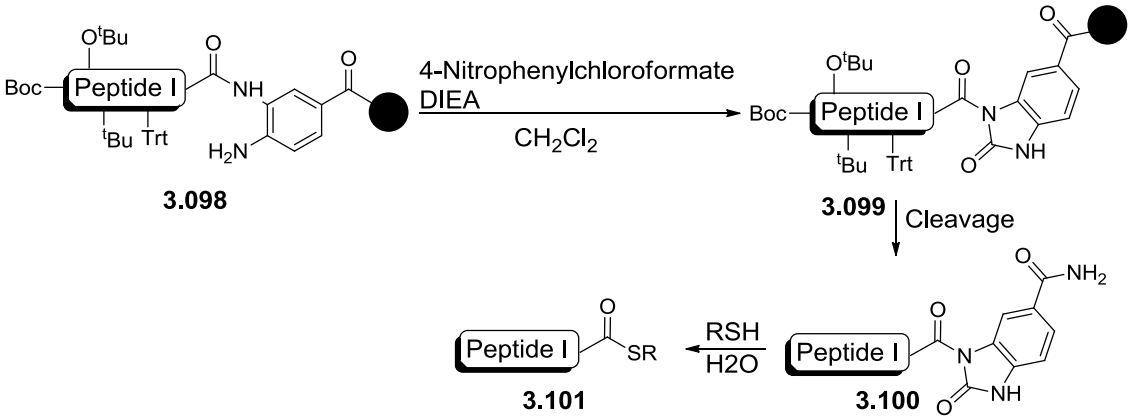

**Scheme 3.18**: General scheme for sequential auxiliary assisted SPCL in a N→C direction by using masked thioesters (**A**) and in a C→N direction relying on orthogonal protecting groups (**B**).

In recent years, much effort was invested in the development of strategies for the synthesis of thioesters following Fmoc procedure. The methods mostly rely on introduction of a masked or latent thioester which is activated post-synthesis to yield a reactive thioester. The unmasking or activation of

the thioester is orthogonal to the synthesis and cleavage conditions used during SPPS and so should be compatible with unprotected peptide fragments. Three approaches were explored: The 3,4-diaminobenzoic acid (Dbz) linker[128], kinetically controlled ligation (KCL)[129] and peptide hydrazides[130, 131].

### 3.2.8.1 Employing 3,4-Diaminobenzoic Acid as Masked Thioester

The method was developed by Dawson for the synthesis of peptide thioesters by Fmoc strategy. The 3,4-diaminobenzoic acid is introduced in the beginning of the sequence and elongation of the peptide is done on one of the amines located on the benzene ring. Acylation of one of the amines deactivates the neighboring one and allows synthesis to be carried out without the need to protect the free benzylic amine. After SPPS is complete the protected Dbz **3.099** is cyclized to form the urea derivative 1H-benzo[d]imidazol-2(3H)-one (Nbz) **3.100**, which possess thioester-like properties and can be substituted to form thioester **3.101** (scheme 3.19). The method was developed for protected peptides as having other free amines on the peptide backbone will lead to unwanted acylation. The MUC1 however, lacks lysine residues and thus only the Dbz amines should be available for acylation which should ensure selective acylation. Any acylation on other residues such as serine or threonine should form labile esters that can be hydrolyzed to retrieve the unmodified residue.




**Scheme 3.19**: Synthesis of peptide thioesters employing Dbz

Aux-MUC1(1-20)-Dbz **3.086** was ligated to resin bound $His_5$-MUC1-SBzl peptide **3.102**. After 16 hr UPLC-MS analysis showed full conversion to the ligation product **3.102** (figure 3.31 B). To form the thioester **3.104** the resin was suspended in DCM and treated with 4-nitrobenzylchloroformate and DIEA. The resin was resuspended in ligation buffer for 10 min to promote thioester formation.

During reaction the beads lost their characteristic blue colour and the texture of the resin changed which suggests modification of the resin's hydroxyl groups. UPLC-MS analysis (figure 3.31 C) revealed no product was formed. The major peak observed was ligation product $His_6$-MUC1(1-20)-Aux-MUC1(1-20)-Dbz-Gly **3.103.** A secondary smaller peak with mass of +167 Da from desired

product **3.104** was also detected. This mass is most likely acylation of the N-terminal hexahistidine chain free amine.

With negative results detected by UPLC as well as the beads incompatibility towards organic solvent the approach was abandoned.

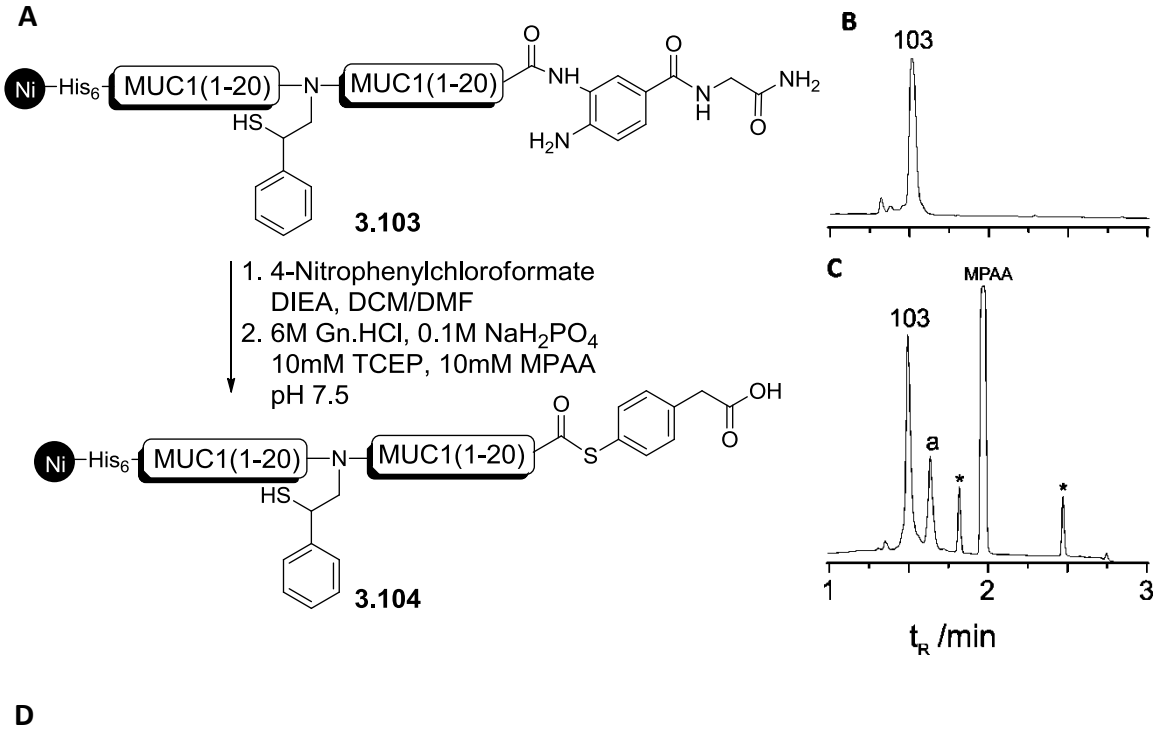
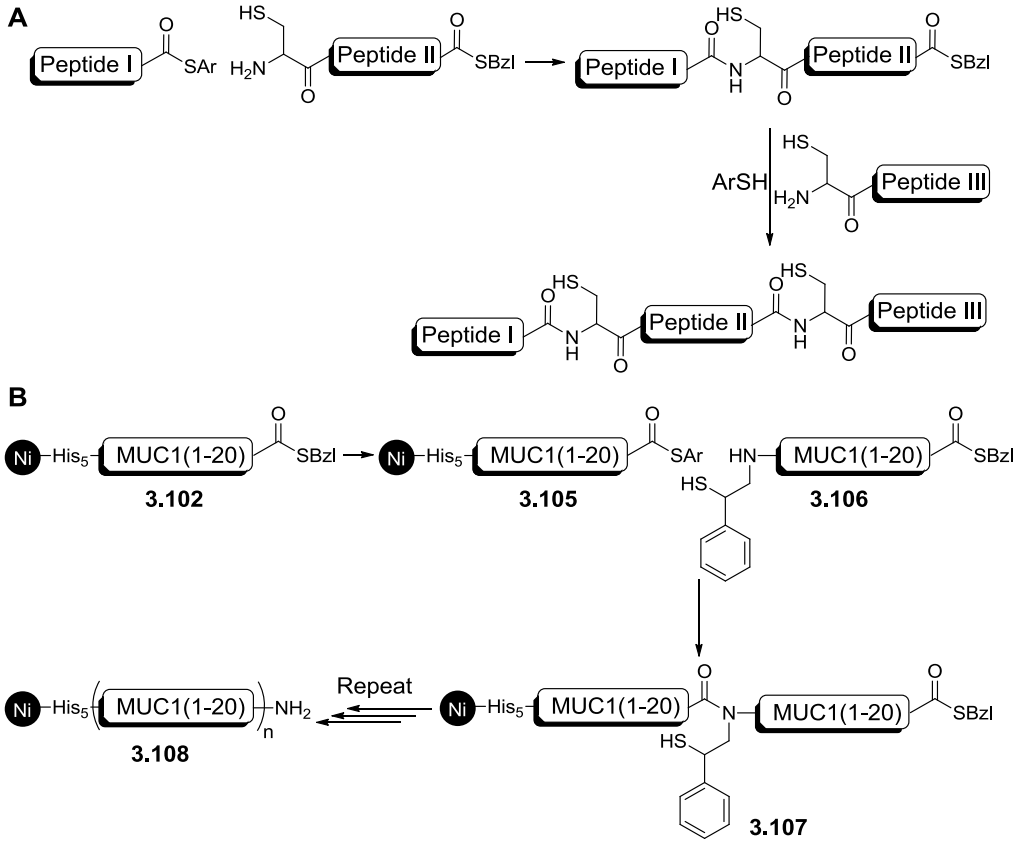



**Figure 3.31**: **A)** Thioester synthesis from Dbz; **B)** UPLC analysis of ligation product $His_6$-MUC1(1-20)-Aux-MUC1(1-20)-Dbz-Gly **3.103**; **C)** UPLC analysis of thioesterification of ligation product **3.103** to give thioester **3.104**; peaks labelled with * correspond to unassigned masses; 03-60% B in 6 min, 210nm.

### 3.2.8.2 Employing Inactive Benzyl Thioesters for Controlled SPCL

Ligation experiments done in solution have shown the auxiliary does not promote ligation when non aromatic thioesters are used. This characteristic of the auxiliary was utilized to try and perform kinetically controlled ligation (scheme 3.20). A benzyl thioester peptide **3.102** anchored to the solid support is activated by thiol exchange with an aromatic thiol (such as MPAA or thiophenol) to yield an active thioester **3.105**. The active thioester is then ligated to a middle fragment **3.106** in absence of thiol additives, ensuring a selective reaction (figure 3.32)




**Scheme 3.20**: **A)** General concept of kinetically controlled ligation; **B)** Controlled auxiliary-assisted SPCL of MUC1.

To form the active thioester the resin bound thioester $His_5$-MUC1(1-20)-SBzl **3.102** was treated with a phosphate buffer containing guanidinium and either MPAA or thiophenol and reactions were monitored by UPLC-MS. In both experiments the major peak detected is the starting benzyl thioester with small peaks corresponding to hydrolyzed thioester (figure 3.32 B and C).

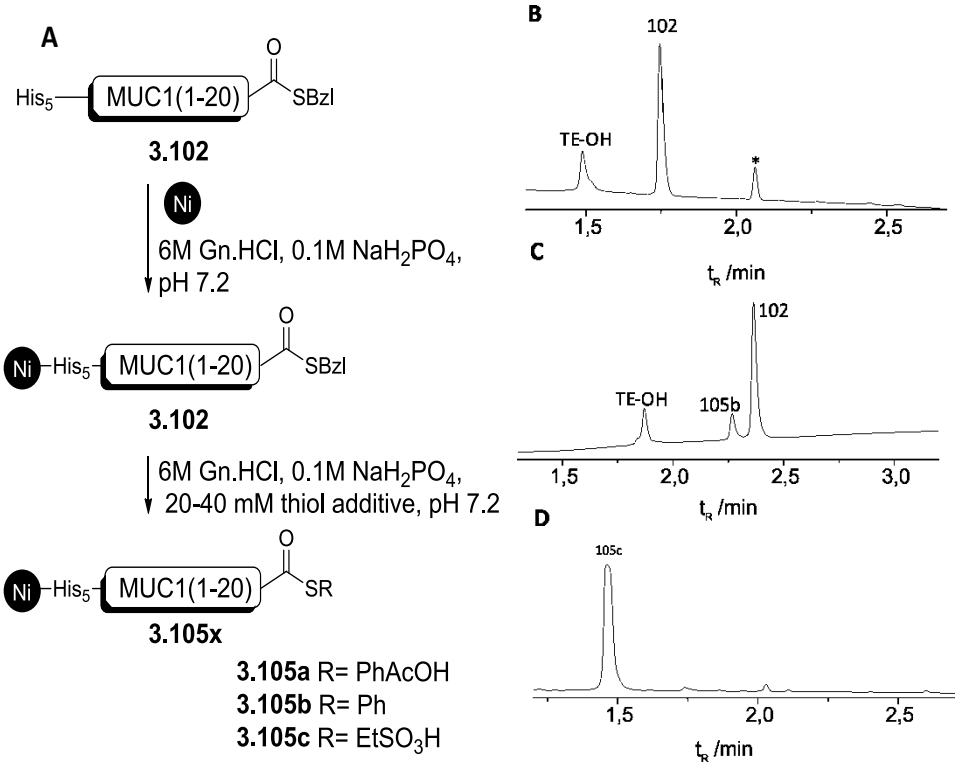



**Figure 3.32**: **A)** Scheme of transthioesterification of peptide **3.121** with different thiols; **B-D)** UPLC analysis of the reaction with MPAA (**B**), thiophenol (**C**) and MESNa (**D**). **TE-OH**: hydrolysed thioester peptide **3.102**; Peak labelled with * is MPAA; Gradient B and D: 03-60% B in 4 min; Gradient C: 03-60% B in 6 min; 210 nm.

A third experiment was performed using MESNa as thiol nucleophile. Being the most active aliphatic thioester [132] it should readily exchange the benzyl thioester and ideally be reactive enough to enable ligation with the auxiliary. Indeed, treating thioester **3.102** in a buffer containing MESNa resulted in full conversion to the MESA thioester **3.105c** (figure 3.32 D).

Next, ligation between the MESA thioester **3.105c** and peptide Aux-MUC1(1-20)-SBzl **3.106** in buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP pH 7.2 was performed. Unfortunately MESA thioester proved to be not reactive enough to go through ligation with only starting material observed by UPLC analysis (figure 3.33). With no ligation taking place with MESA thioester and activation with more reactive aromatic thiols not proceeding the approach was terminated.

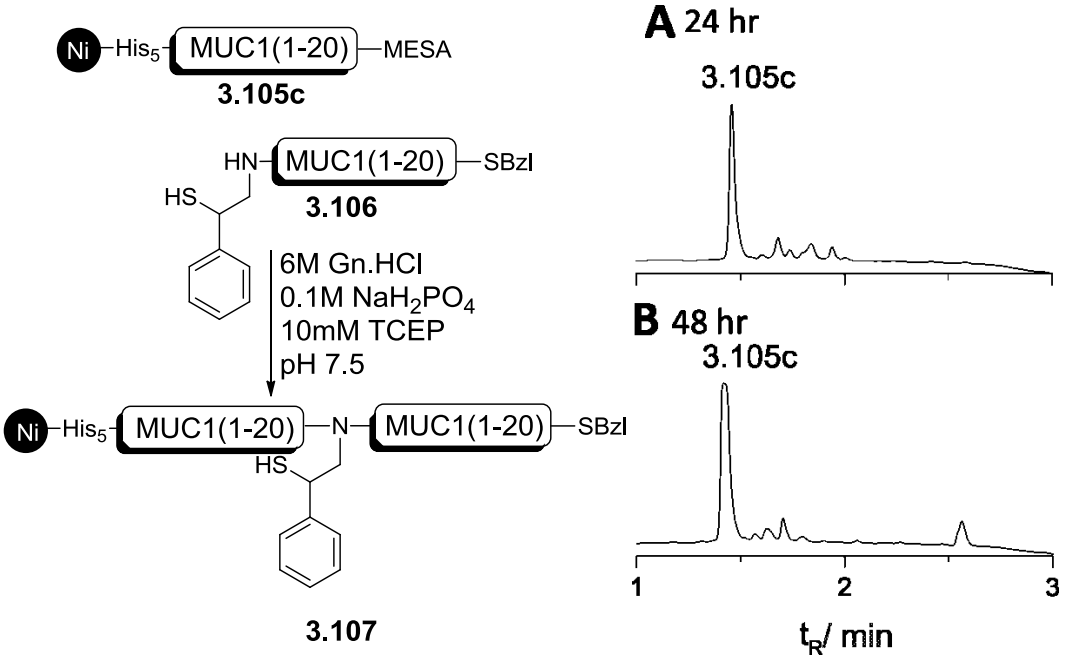
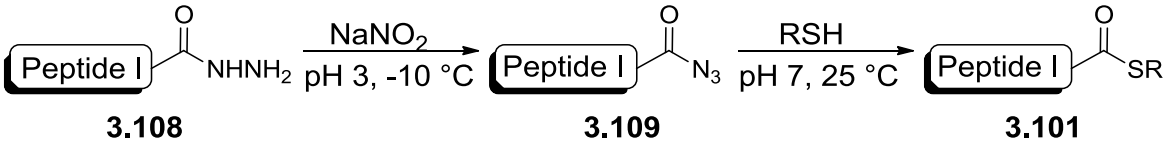



**Figure 3.33**: UPLC analysis of ligation between peptide **3.105c** and **3.106** at **A)** 24 hr **B)** 48 hr; 03-60% B in 4 min, 210 nm.

### 3.2.8.3 Employing Peptide Hydrazides

Similar to Dbz, peptide hydrazides also function as thioester precursors and can be transformed to thioesters. By treating the unprotected peptide hydrazides **3.108** with aqueous solution of sodium nitrite at slightly acidic pH, peptide azide **3.109** is formed. The azide in turn is being substituted to a thioester via a nucleophilic attack by a thiol (scheme 3.21).




**Scheme 3.21**: General concept of thioester formation employing peptide hydsrazides.

The peptide hydrazide was synthesized as described in the literature [130]. Chlorotrityl resin was treated with hydrazine in the presence of $NEt_3$ in DMF to yield resin **3.110**. The sequence was completed by automated SPPS. HPLC purification gave $His_6$-MUC1-$NHNH_2$ peptide **3.111** with 55% isolated yield (figure 3.34).

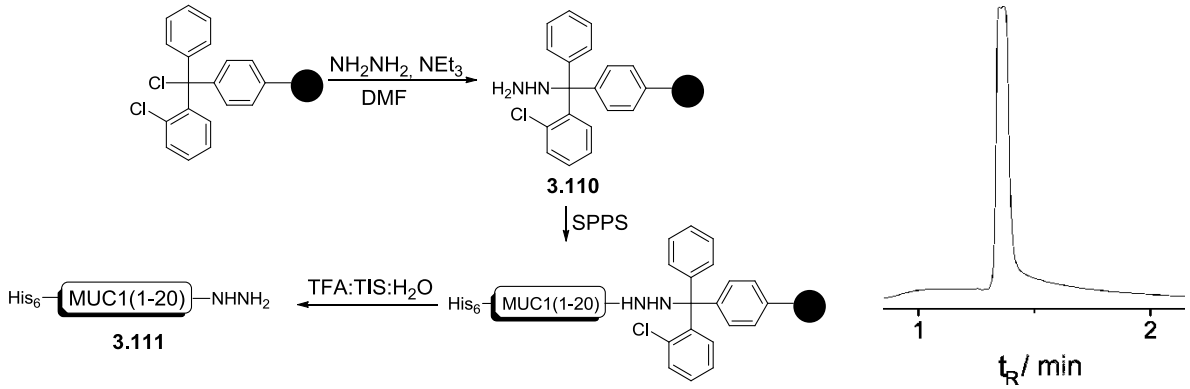



**Figure 3.34**: UPLC analysis of $His_6$-MUC1(1-20)-$NHNH_2$ **3.111**; Gradient: 03-60% B in 4 min; 210 nm.

The peptide **3.111** was anchored to the Ni-NTA Resin employing the previously described method and suspended in aqueous solution of 30 mM $NaNO_2$ pH 3 at -10 °C for 30 min followed by addition of a phosphate buffer containing 20 mM MPAA pH 7.2 (final pH value is 6.8) at 25 °C for 30 min. UPLC-MS analysis showed no reaction took place with a single peak corresponding to the hydrazide peptide **3.111** observed. Longer exposure times, increasing concentration of $NaNO_2$ and MPAA or performing the two steps separately by first washing the resin from the nitrite buffer and then adding the MPAA containing buffer all did not change the result (figure 3.34 A). As the experiments in the previous section demonstrated thioesterification with MPAA is problematic the experiments were repeated using MESNa as thiol. Exchanging the thiol had no influence and no reaction was detected (figure 3.35 B).

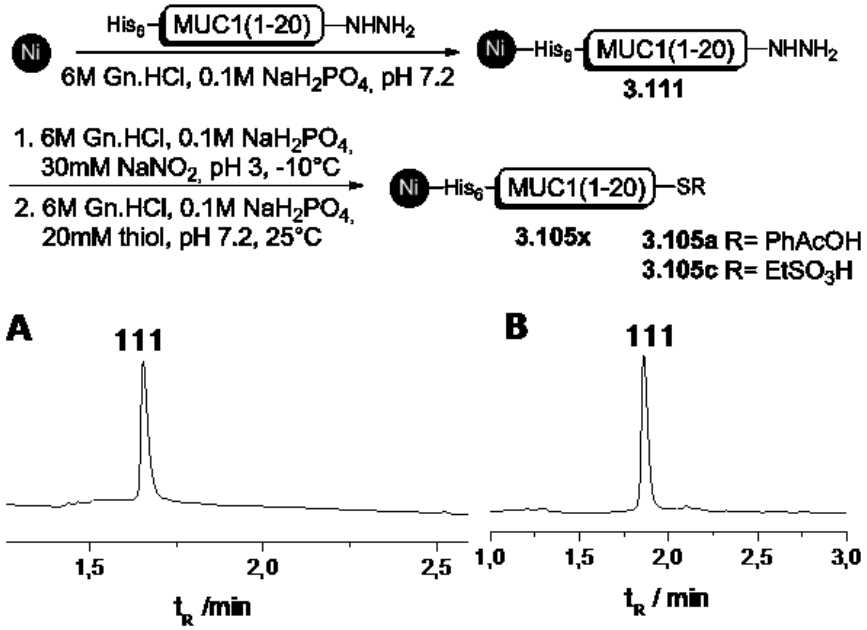



**Figure 3.35**: UPLC analysis of peptide hydrazide **3.111** transformation to thioester. **3.105**. **A)** with MPAA; **B**) with MESNa; 03-60% B in 4 min, 210nm.

As control, the reaction was performed in solution using the same conditions. In the control experiment thioester formation was detected (figure 3.36) meaning that the problem does not lie with the system itself but rather stems from some kind of hindrance presented by the solid-phase. Without being able to form the thioester the route was abandoned as well.

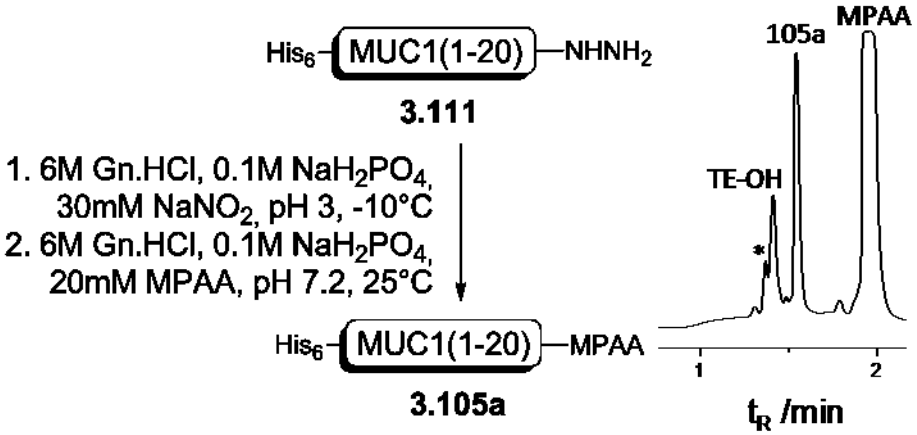



**Figure 3.36**: UPLC analysis of peptide hydrazide **3.111** transformation to thioester. **3.105a**; TE-OH: hydrolysis of thioester 3.100; Peak labelled with * correspond to unassigned masses; 03-60% B in 4 min, 210nm.

### 3.2.8.4 Efforts toward Sequential Ligations: Summary

In this section an attempt to perform sequential ligation on solid support was made. Utilizing latent thioester or thioester precursors to achieve selective ligation between two peptide fragments for the selective chemical ligation between a thioester H-MUC1(1-20)-SR and a middle fragment Aux-MUC1(1-20)-X (X is an inactive crypto thioester or low reactive thioester). Three approaches were tested: The N-acyl urea approach, formation of highly active thioesters pre-ligation and performing KCL and in-situ conversion of peptide hydrazide into peptide thioester.

In the first case, a peptide was synthesized on a Dbz linker to give the Aux-MUC1(1-20)-Dbz-Gly **3.085**. The peptide was successfully ligated to a resin anchored $His_6$-MUC1(1-20)-SBzl **3.093** under standard ligation conditions to yield peptide $His_6$-MUC1(1-20)-Aux-MUC1(1-20)-Dbz-Gly **3.095**. Activation of the Dbz moiety to the active Nbz species was tried following the known protocol [128] which is done in organic medium of DMF and DCM. These conditions proved incompatible with the resin leading to its decomposition.

Use of differently active thioester peptides to perform KCL was tried next. By forming an active aryl thioester prior to adding the middle fragment and performing ligation in absence of thiol additive a selective ligation will take place leading to formation of a single ligation product. Peptide thioester $His_6$-MUC1(1-20)-SBzl **3.093** was anchored to Ni-NTA resin and treated with aqueous buffer containing MPAA, thiophenol or MESNa. Treatment with MPAA or thiophenol did not yield

the desired respective thioesters $His_6$-MUC1(1-20)-MPAA **3.105a** or $His_6$-MUC1(1-20)-SPh **3.105b** and starting benzyl thioester **3.093** remained as major species. Conversion to the MESA thioester **3.105c** was quantitative however, the thioester proved too unreactive to participate in the ligation and no product was detected. While in-situ formation of MPAA or thiophenol thioesters in the presence of a cysteinyl or cysteinyl equivalent peptide is possible, as the active thioester is immediately attacked by the cysteinyl peptide to forms the stable thioester or amide bond, isolation of such active thioesters is challenging due to the very high reactivity of such species. Similar results were encountered on other occasions when attempts to prepare MPAA thioesters in solution by transthioesterification reached a conversion lower than 50% despite using high excess of reactants or prolonging the reactions time.

Peptide hydrazides have previously been used as masked thioesters [133-135]. Classic procedure to convert these species into active thioesters required protection of other amine and thiols on the sequence and working in organic solvents. Recently, Liu et al. have developed a new procedure that allows peptide thioesters to be formed in-situ from peptide hydrazides [130] and have also demonstrated the peptide hydrazide utility in KCL [131]. While conversion to the thioesters proceeded well in solution, the peptide hydrazide was completely inert once anchored to the Ni resin. The reason behind the hydrazide inertness is the strong affinity of the hydrazide moiety towards the Ni(II) ions [136] Similar to the imidazole, the hydrazine moiety forms a highly stable complex with the Nickel which renders it unavailable and completely hinders the reaction.

## 3.2.9 Auxiliary-Assisted SPCL of a 80mer VNTR MUC1 Peptide

The experiments on the nickel solid support have shown that the support is applicable for the ligation chemistry being water swellable and stable towards the reagents used in ligation. Although attempts to perform sequential ligation from the N→C direction did not yield a positive result, ligations of two fragments were easily accessible. The MUC1 synthesis has also been optimized throughout the experiments reaching highly effective synthesis with good yields and minimum truncation sequences. Through this optimization process, the 40mer MUC1 peptide, consisting of two tandem repeat sequence, was synthesized as well. As a final goal and proof of applicability of the auxiliary assisted SPCL, the 80 amino acids long MUC1 VNTR peptide was to be synthesized employing a single ligation step. Ligation of longer peptide fragments is often slower and leads to a larger amount of side products due to incomplete ligation. It is in cases such as these where the solid support internal purification characteristic makes it an especially attractive technique. As any excess reagents and unreacted reactants can be washed off the resin, it should give a clean ligation profile and easy isolation of the product.

### 3.2.9.1 Synthesis of Fragments MUC1(1-40)-SBzl 3.112 and Aux-MUC1(41-80)-$His_6$ 3.114

The 40mer benzyl thioester MUC1(1-40)-SBzl **3.112** was synthesized as described in the previous section employing the safety catch linker. After purification on a preparative HPLC thioester **3.112** was isolated with 35% yield (figure 3.37).

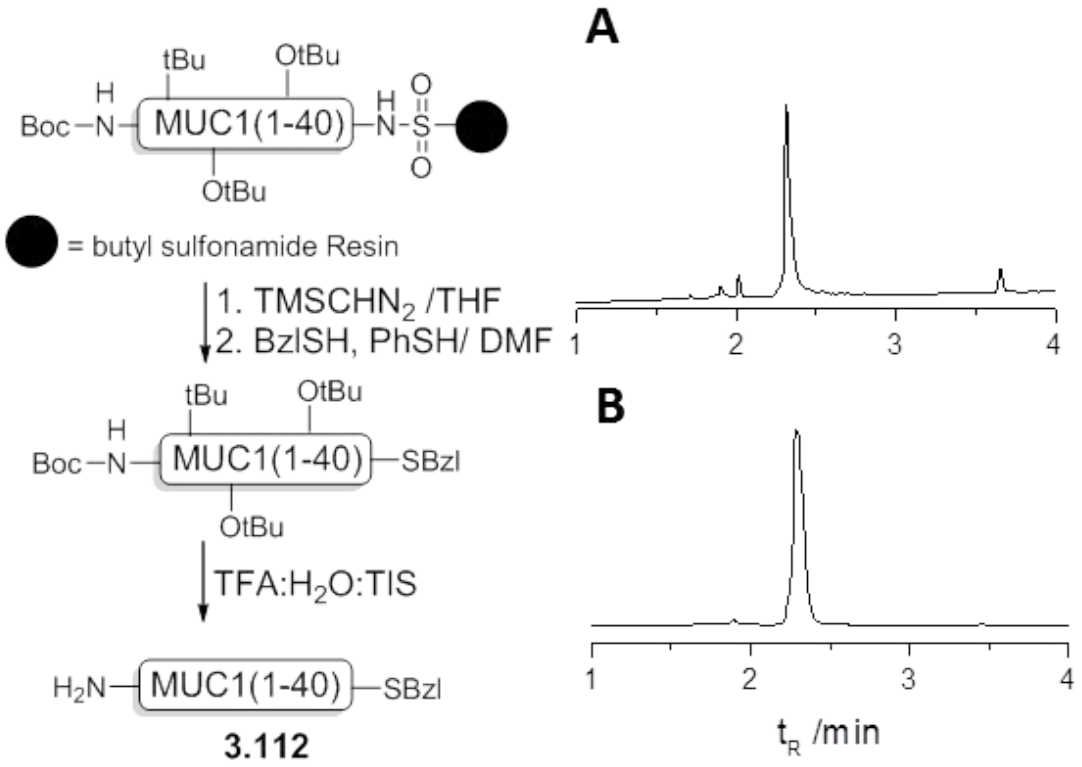



**Figure 3.37**: UPLC analysis of crude (**A**) and pure (**B**) H-MUC1(1-40)-SBzl thioester peptide **3.112**; 03-60% B in 6 min, 210 nm.

The auxiliary peptide Aux-MUC1(41-80)-$His_6$ **3.114** was synthesized following standard SPPS procedure. After synthesis was complete the auxiliary was attached to the peptide using reductive amination. The peptides were cleaved and purified on preparative HPLC to yield the pure peptides with 15% yield (figure 3.38).

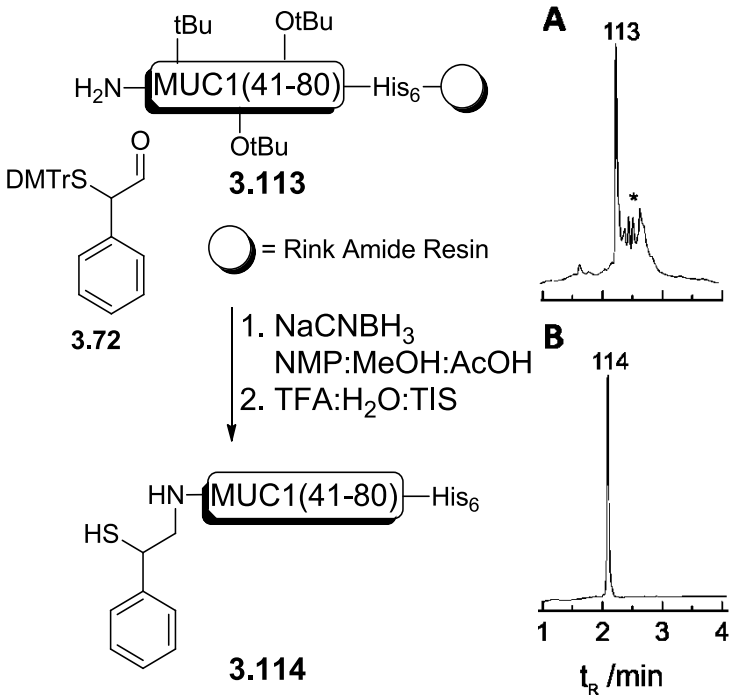



**Figure 3.38**: UPLC of Aux-MUC1(1-40)-$His_6$ **3.113**. **A)** Crude H-MUC1(1-40)-$His_6$ **X**; **B)** Pure peptide Aux-MUC1(1-40)-$His_6$ **3.114**; 03-60% B in 6 min, 210 nm.

### 3.2.9.2 Auxiliary Assisted SPCL towards MUC1(1-80) VNTR

With the two peptide fragments isolated, auxiliary-assisted SPCL was performed as follows: auxiliary peptide **3.114** was anchored to the nickel support as previously described. Complete attachment was observed after 1.5 hr. The resin was washed and the thioester peptide **3.112** (1.2 eq), dissolved in a buffer of 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 10 mM TCEP, 20 mM MPAA, pH 7.5, was added to resin. The reaction was performed at 25 °C and monitored by UPLC-MS. A fresh batch of thioester **3.112** (0.5 eq) was added in 24 hr intervals. The process was repeated two times after which complete conversion to the desired ligation product **115** was detected by UPLC-MS analysis (figure 3.39).

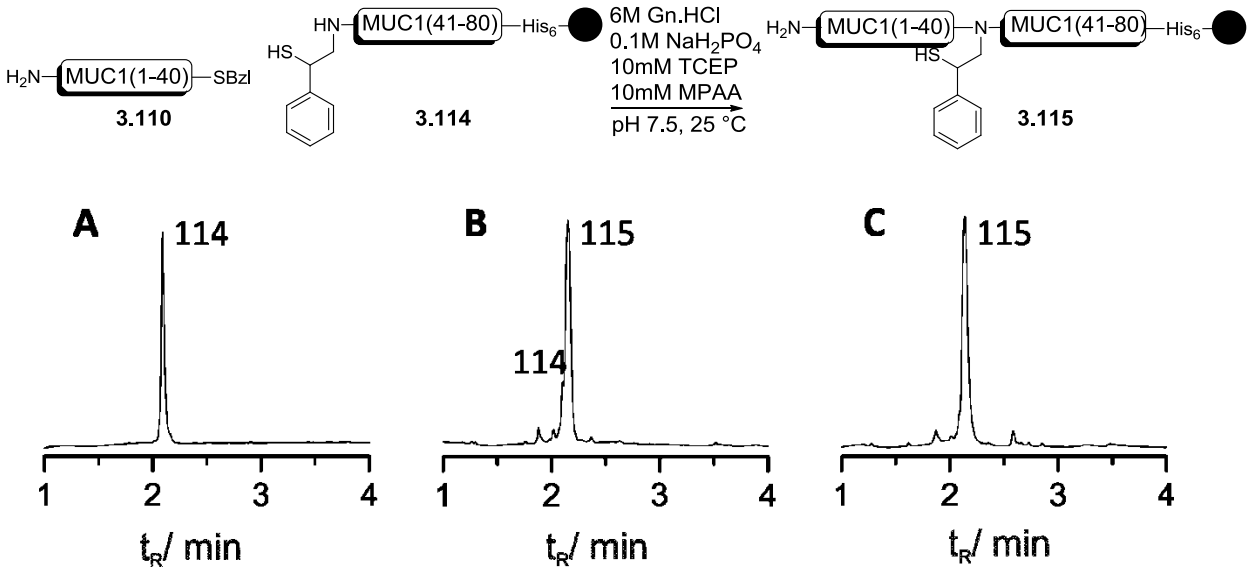



**Figure 3.39**: UPLC analysis of SPCL between H-MUC1(1-40)-SBzl **3.112** and Aux-MUC1(1-40)-$His_6$ **3.114** to give ligation product **3.115** at **A)** 0 hr; **B)** 24 hr; **C)** 72 hr; 03-60% B in 6 min, 210 nm.

### 3.2.9.3 Removal of Auxiliary from Ligation Product MUC1(1-80) VNTR

In order to remove the auxiliary 20 µl of an aqueous buffer containing 4 M morpholine and 1 M TCEP (pH 8.5) was added to 20 µl elution solution containing the released ligation product **3.115** for 16 hr at 25 °C (figure 3.40). The product was directly isolated by semi-preparative HPLC to yield the desired H-MUC1(1-80)-$His_6$ **3.116** in 45% yield (figure 3.40).

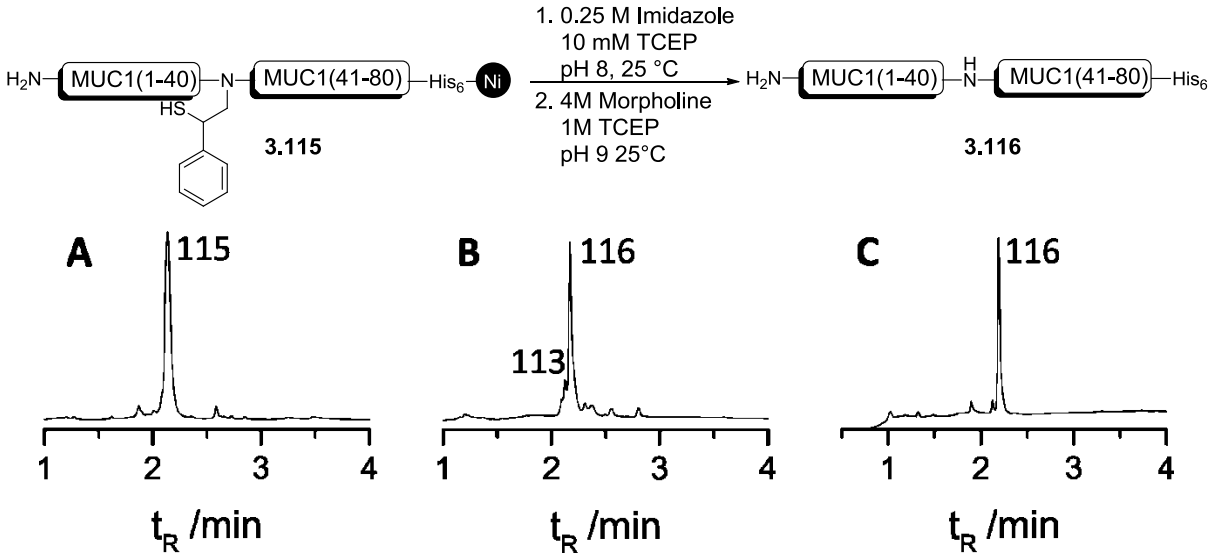



**Figure 3.40**: UPLC analysis of removal of auxiliary from crude ligation product **3.115** to give H-MUC1(1-80)-$His_6$ **3.116**; **A)** 0 hr; **B)** 16 hr; **C)** Purified H-MUC1(1-80)-$His_6$ **3.116**; 03-60%B in 6 min, 210nm.

### 3.2.10 Auxiliary-mediated SPCL Conclusions

To date, most proteins and peptides synthesized by NCL and its numerous extensions are synthesized in solution. Working in solution allows the reaction to be directly and swiftly set up, easily analyzed and purified without the need for work-up step. Solution phase NCL is indispensable and advancements in performing multiple ligations sequentially are constantly being made. Despite the big leaps the field has gone through, sequential NCL is still limited. Incompatible reagents prevent continuation of the protein chain, thiol additives prevent desulfurization, protecting groups such as acetamidomethyl (Acm) require different medium for removal, or unmasking of thioesters becomes incompatible with other functional groups. Furthermore, ligation of several fragments leads to accumulation of side products and unreacted reactants.

As the field of chemical synthesis matures and the complexity of desired targets increases attempts are continuing ever more often to adapt NCL to be performed on solid phase. The motivation behind these trials is to exert the advantages SPPS gave the field of peptide synthesis and lend them to the field of total chemical synthesis of proteins, i.e. ability to use large excess of reactants to force the ligation to completion, performing consecutive reaction in one-pot, easy removal of side-product and excess reagents by simple washing steps and the possibility to automatize the process. SPCL has proven itself as a very powerful tool for sequential ligation. In this work, efforts towards the first auxiliary mediated SPCL has been taken and successfully achieved.

Appropriate solid-support was found in the form of the His-Tag affinity resin. It has excellent swelling properties in aqueous buffers, it is durable against thiols and phosphines and provides orthogonal anchoring and release of the peptides onto and from the support.

The auxiliary presented in this work has allowed the first successful auxiliary-mediated SPCL. Ligations of peptides fragments 20 residues in length were complete within 15 hr with exceptional purity. Ligation of 40mer fragments, which are often slow even when performed in solution, were successful as well, reaching completion in 3 days thus further highlighting the usefulness and usage of this new auxiliary.

The ability to remove excess reagents and reactants by washing allowed high equivalents of thioester to be introduced while avoiding troublesome purification. The mild conditions required to release the peptides from support proved compatible with the conditions required for removal of the auxiliary, further minimizing purification steps and allowing the entire synthesis to proceed as a continuous process.

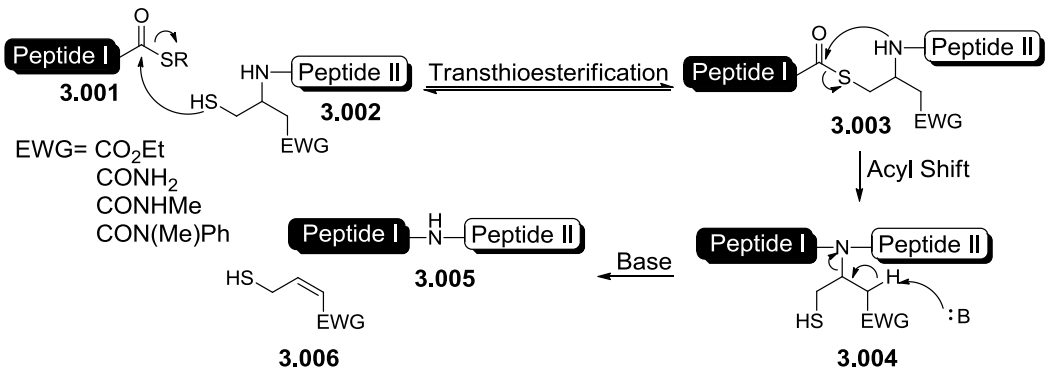
To summarize, the first auxiliary assisted SPCL was demonstrated utilizing the newly conceived 2-mercapto-2-phenylethyl auxiliary **3.073**. The auxiliary was effectively and rapidly coupled to the peptide directly during SPPS and promoted SPCL at peptide fragments of various lengths effectively and cleanly. The mild removal chemistry of the auxiliary as well as the mild release of the product

from the solid support makes the method ideal for synthesis of peptides with acid sensitive post translational modification such as phosphorylation, sulfonation or glycosylation.

# 4 Summary and Outlook

## 4.1 Summary

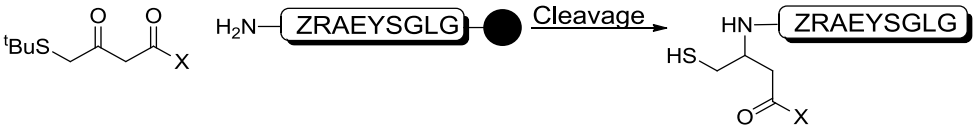
$N^{\alpha}$-auxiliaries for assisting NCL have been a topic for research since very early on after the introduction of NCL. With the potential to resolve the requirement for a cysteine residue at the ligation site, which continues to be the major limitation of NCL, and potentially allow ligation to be performed in a variety of junctions, they stood in the center of several researches. Despite advancements and progress taken in the field over the years, the auxiliaries have yet to fulfill their full potential, mostly due to inability to assist ligation at difficult junctions. In this study, the scope of auxiliary assisted NCL was extended. A novel family of $N^{\alpha}$-auxiliaries was designed, developed and utilized for the synthesis of peptides and proteins. The auxiliaries aided the chemoselective condensation of peptide fragments to form full-length peptides and proteins of different length in targets unattainable by conventional NCL thus extending the chemists available tools for synthesis of proteins. In the first part of the dissertation a group of auxiliaries were designed to be labile towards mildly basic conditions by abstraction of an acidic proton located on the auxiliary, leading to cleavage from the peptide via an elimination mechanism. Four auxiliaries, based on 4-mercaptobutyrate scaffold, were synthesized, each featuring a different electron withdrawing group. The auxiliaries were studied for their ability to mediate native chemical ligation (scheme 4.1).




**Scheme 4.1**: Structure and proposed mechanism of auxiliary-assisted NCL and base induced cleavage of auxiliaries to yield the native peptide bond.

The auxiliary peptides were assembled directly on the solid support by reductive amination between the N-terminal amine of the peptide and a ketone function located on the auxiliary precursor backbone. This synthetic route gives universality to the method as the auxiliaries can be coupled to any peptide without the need to synthesize modified amino acid building blocks for different ligation junctions (scheme 4.1).

**Table 4.1: Yields of auxiliary peptides 3.42-3.49.**




**3.021** X= OEt
**3.024** X= $NH_2$
**3.028** X= NHPh
**3.031** X= NMePh

**3.032** Z =G
**3.033** Z= A

**3.042** Z= G X= OEt
**3.043** Z= A X= OEt
**3.044** Z= G X= $NH_2$
**3.045** Z= A X= $NH_2$

**3.046** Z= G X= NHPh
**3.047** Z= A X= NHPh
**3.048** Z= G X= NMePh
**3.049** Z= A X= NMePh

| Auxiliary Peptide | Isolated yield | Auxiliary Peptide | Isolated yield |
|---|---|---|---|
| **3.042** | 30% | **3.046** | 36% |
| **3.043** | 14% | **3.047** | 15% |
| **3.044** | 24% | **3.048** | 16% |
| **3.045** | 18% | **3.049** | 15% |

The efficiency of the four auxiliaries in mediating ligation was tested in model studies at four junctions, each with increasing sterical demand: Gly-Gly, Ala-Gly, Gly-Ala and Ala-Ala. Product formation over time for each of the auxiliaries is shown in figure 4.1. The data clearly shows that ligation is highly effective at Gly-Gly junction but efficiency quickly drops with increase in sterical demand. Ala-Ala junction draws the limit of the auxiliaries' ability to mediate ligation, with no ligation product formed.

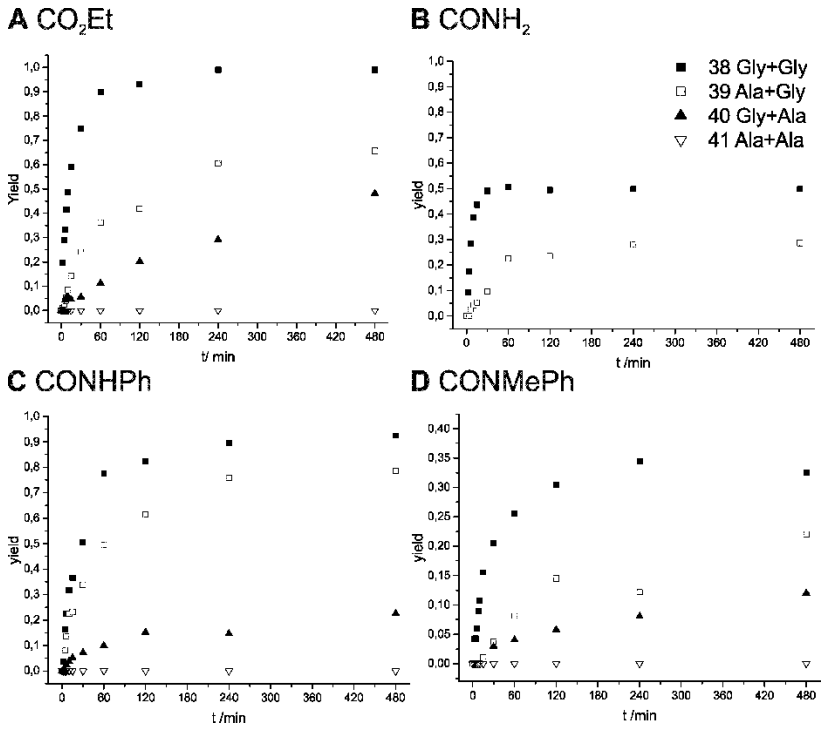



**Figure 4.1**: Product formation over time of the four auxiliaries at different ligation junction.

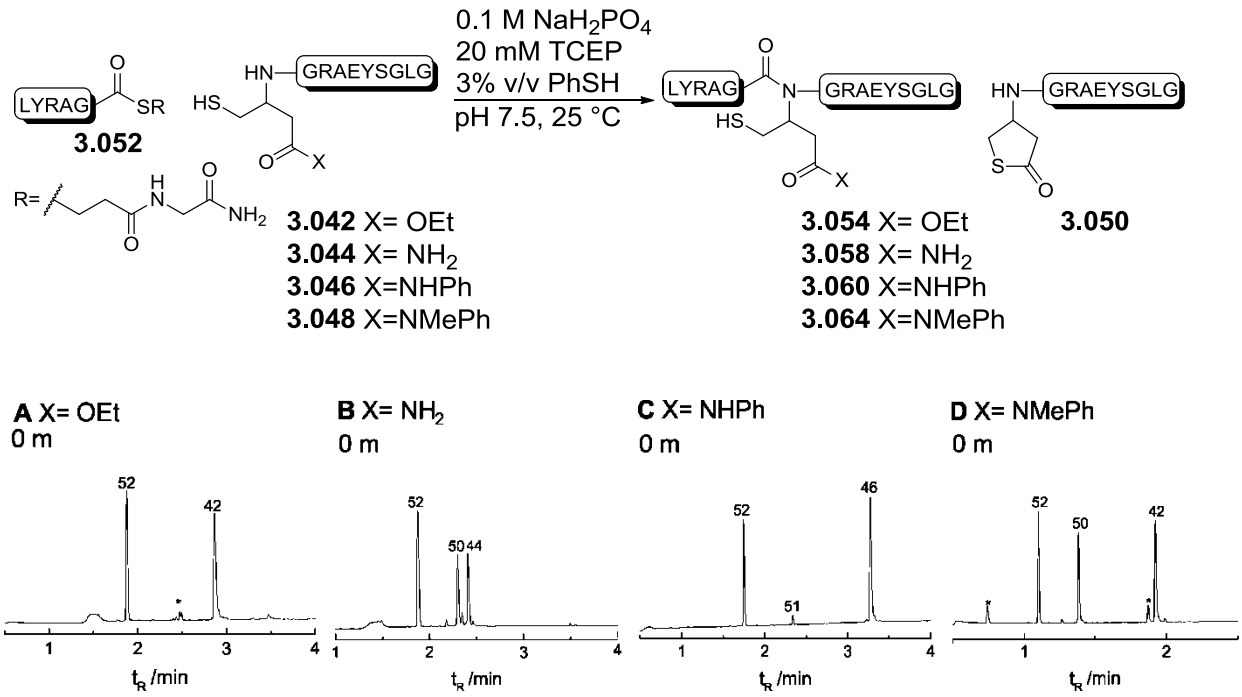



**Figure 4.2**: UPLC analysis of measurements at 0 hr of ligations between thioester **3.052** and auxiliary peptides **3.042** (**A**), **3.044** (**B**), **3.46** (**C**) and **3.048** (**D**) demonstrating auxiliaries' tendency to form the thiolactone species **3.050**; 03-60% B in 6 min, 210 nm.

Between the four auxiliaries, it was the ethyl ester auxiliary **3.016** and N-phenylamide auxiliary **3.018** that have shown the best results while the amide auxiliary **3.017** and N-methyl-N-phenyl auxiliary **3.019** gave poor results throughout the entire study. The large difference between the auxiliaries does not stem solely from the differences in the auxiliaries' rate constants which only differ slightly from one auxiliary to the next, but from a second property all the auxiliaries possess: formation of a 5- membered thiolactone ring via cyclization. Once in this structure, the auxiliaries cannot mediate the ligation as the thiol moiety is no longer available. The thiolactone form itself is highly susceptible and is swiftly being removed from the peptide backbone to yield the free amine peptides GRAEYSGLG **3.032** or ARAEYSGLG **3.033** as seen when peptide thiolactone **3.050** was dissolved in ligation buffer to yield the unmodified peptide **3.032** within 2 hr (figure 3.6 B).

Despite differences in the tendency to form the thiolactone, all auxiliaries formed it over time. In ligation sites with increased sterical hindrance, when the auxiliary assisted ligation could not proceed (as is the case for the Ala-Ala junction), the formation of the thiolactone and its consequential removal to give peptides **3.032** or **3.033** became the major path (figure 4.3).

The removal of the thiolactone species under ligation condition was only observed on the auxiliary peptide. Once ligation product forms no removal of auxiliary was detected. This behavior seems to contradict known principles about leaving group tendency. Amides are much better leaving groups than amines due to the stabilization of the nitrogen lone-pair electrons along

the amide bond, reducing their basicity. This opposite behavior of the system hints towards auxiliary removal via a radical pathway.

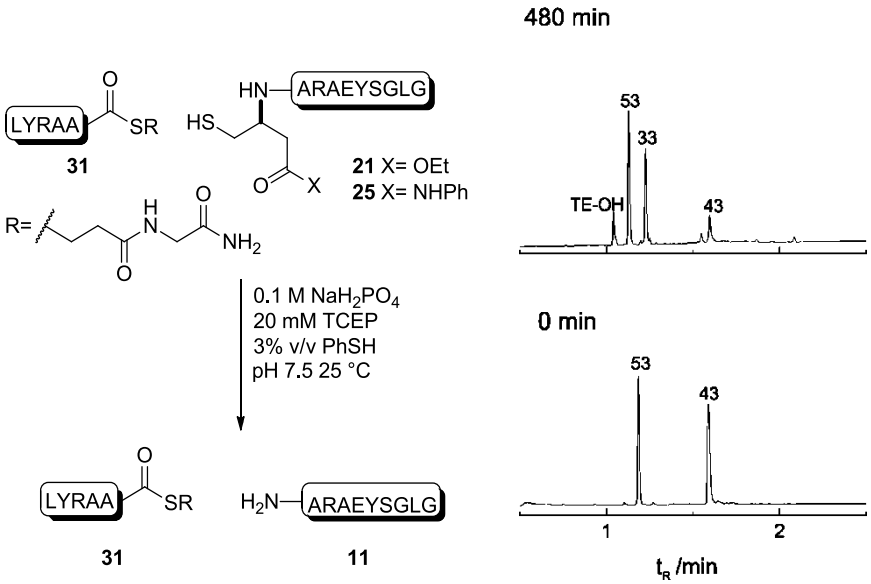



**Figure 4.3**: UPLC analysis of ethyl ester auxiliary assisted NCL at Ala-Ala junction at 0 and 24 hr; **21**: Aux-ARAEYSGLG, **31**: LYRAA-SR, **11**: ARAEYSGLG, **TE-OH**: LYRAA-OH; 03-60% B in 6 min, 210 nm.

The results of the model studies showed only ethyl ester auxiliary **3.016** and N-phenylamide auxiliary **3.018** were stable enough to be used in auxiliary-assisted NCL. As ethyl ester auxiliary **3.016** shown better results and is more readily synthesized, the three other auxiliaries **3.017**, **3.018**, and **3.019** were discarded from further experiments.

To cleave the auxiliary from ligation product aqueous buffer containing TCEP and one of the bases: NaOH, morpholine, piperidine or EDMA were tested. Between the four bases only morpholine and piperidine have cleaved the auxiliary and gave the native, unmodified ligation product (figure 4.4A). Examining the efficiency of the two bases showed that the more basic piperidine was more potent but produced more unidentified impurities. Morpholine produced a cleaner reaction with fewer side products while still exhibiting satisfactory removal rates (figure 4.4B and C).

With experimental data validating the applicability and the initial theorem of the base labile auxiliary on short model peptides, a more challenging target was chosen to test the auxiliary ability to promote ligation on complex, biological relevant systems.

As a target the antimicrobial peptide DCD-1L was chosen. The 48 residues long peptide was split into two fragments with ligation site at Gly-Gly. After 7 hr full conversion to ligation

product was detected. Ligation product was purified by semi-preparative HPLC in 50% yield. The pure peptide was subjected to removal buffer for 6 hr to give the native DCD-1L peptide isolated in 40% yield (figure 4.5).

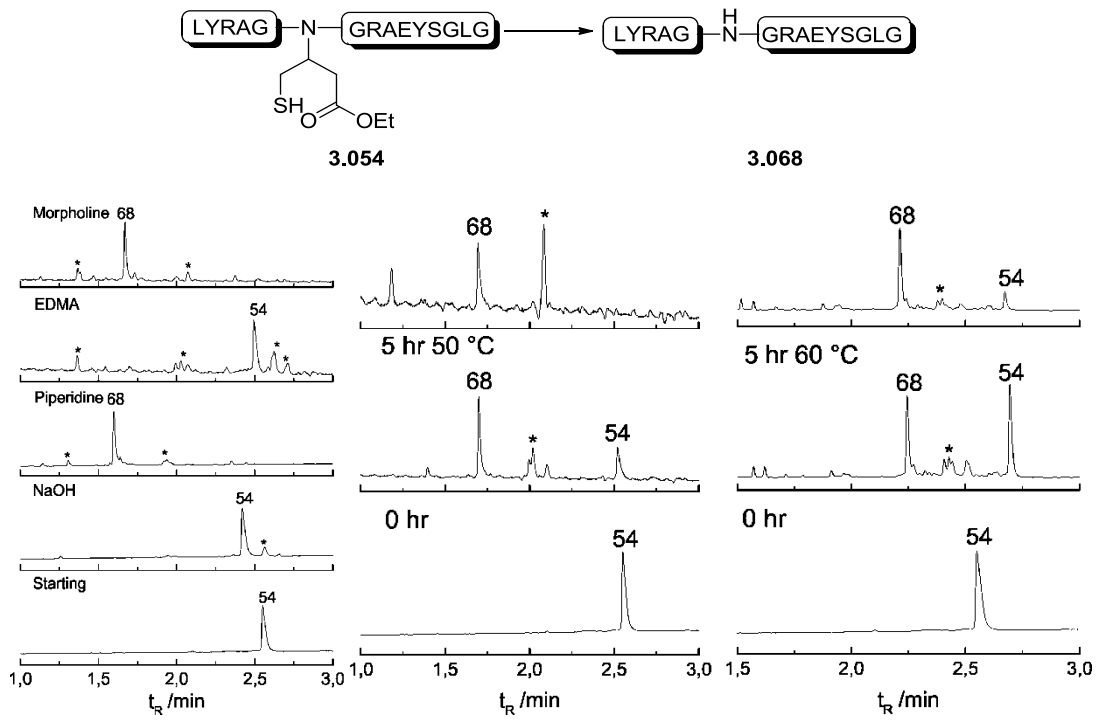



**Figure 4.4**: UPLC analysis of removal of auxiliary from ligation product **3.054** to give native peptide **3.068** after 5 hr. **A)** Screening of base component; **B)** Piperidine treatment at 50 °C and 60 °C; **C)** Morpholine treatment at 50 °C and 60 °C; 03-60% B in 6 min, 210 nm.

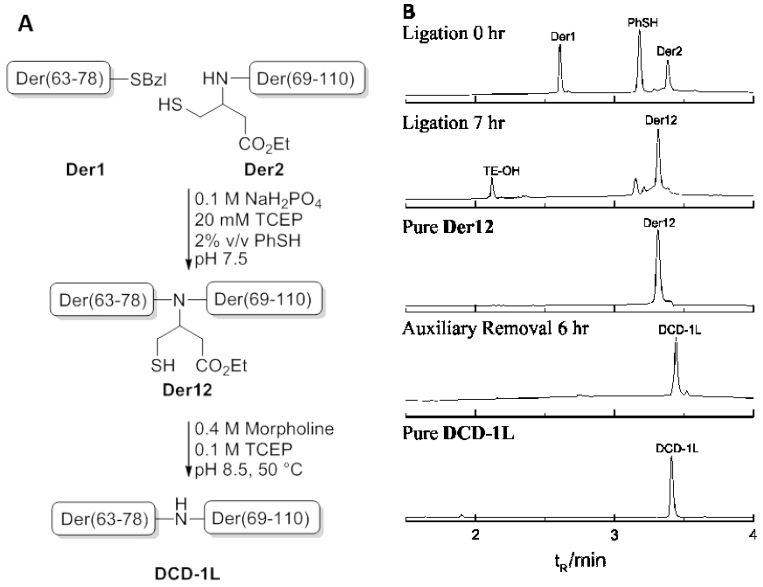
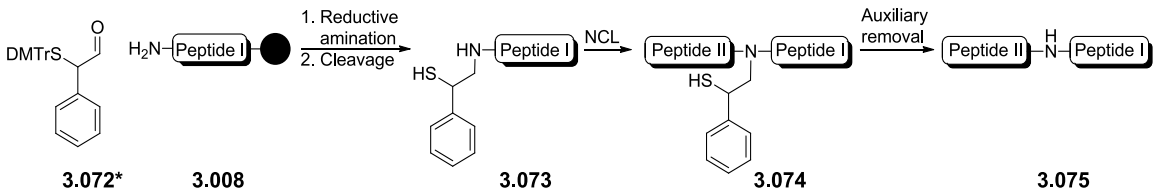



**Figure 4.5**: Synthesis of DCD-1L. **A)** Scheme depicting synthetic route; **B)** UPLC analysis of ligation between **Der1** and **Der2** and auxiliary removal to yield native **DCD-1L**; **C)** MS analysis **Der12** with observed mass 4963.6 Da (calc. 4963.7 Da); 03-70% B in 6 min, 210 nm.

In the second part of the work, a different class of $N^{\alpha}$-auxiliary, 2-mercapto-2-phenylethyl, developed by Loibl et al. [95] was applied for auxiliary assisted solid phase chemical ligation for synthesis of MUC1 tandem repeat domain. In comparison to the previous auxiliaries the auxiliary exhibits better reactivity under reductive amination condition thanks to an aldehyde functionality instead of a ketone, complete stability under ligation conditions and ability to mediate ligation in a wide range of sites (scheme 4.2).




**Scheme 4.2**: General concept for synthesis of auxiliary peptides, ligation and auxiliary removal.

The auxiliary assisted SPCL was performed on water swellable His-Tag resin to enable easy, mild and orthogonal binding and release of the peptide fragments onto- and off the resin. The ligations were tested in both C→N and N→C directions with both directions giving excellent results. The ligation product was released from support by imidazole treatment and auxiliary was removed in one-pot by adding the removal buffer (figure 4.6).

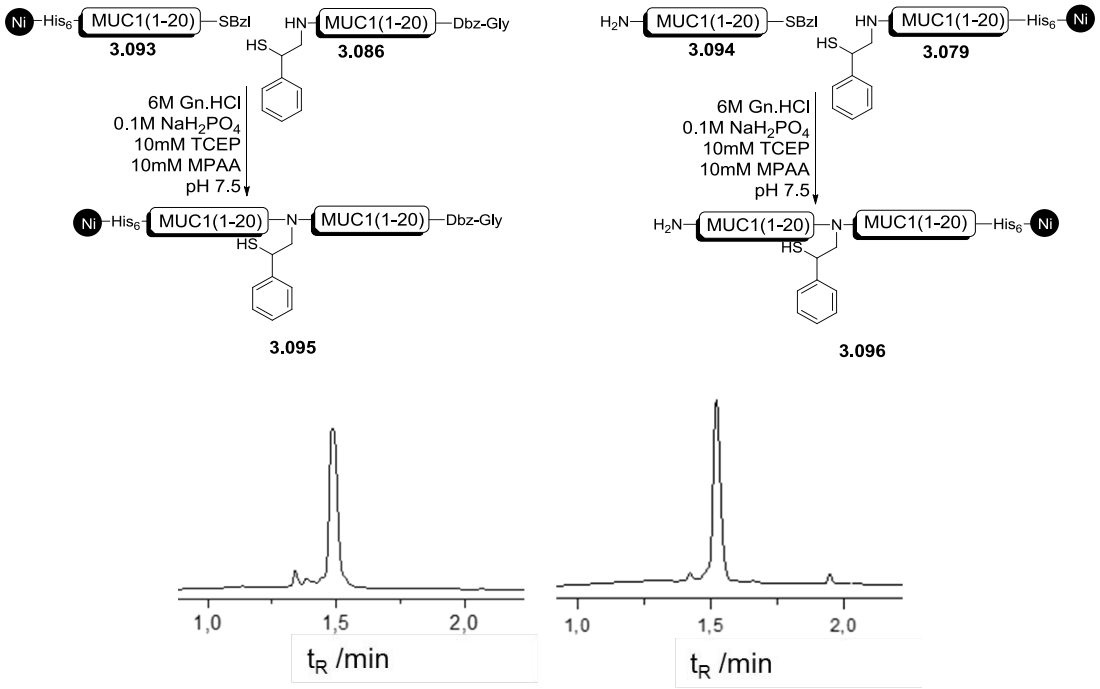
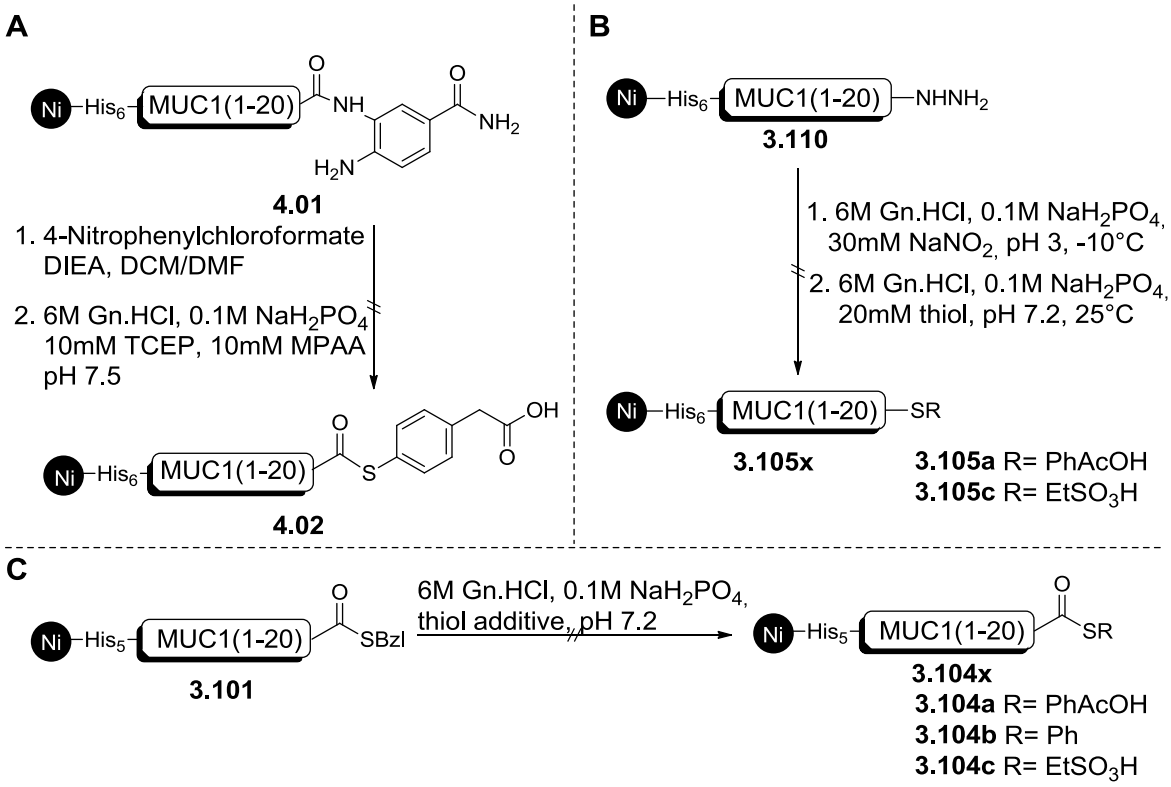



**Figure 4.6**: UPLC analysis of auxiliary assisted SPCL. **A)** N→C directed; **B)** C→N directed; 03-60% B in 6 min, 210 nm.

Efforts towards sequential SPCL were taken as well. Concentrating on N→C directed ligation, three strategies relying on employing masked or hindered thioesters were tested (scheme 4.3): 3,4-diaminobenzoic acid (Dbz) linker used in Fmoc synthesis of thioesters is activated post synthesis by cyclization to form the urea derivative which can be substituted to give a thioester. The method proved incompatible with the polymer support that does not tolerate organic solvents that are required to activate the precursor. A second approach involved peptide hydrazides. These are synthesized in a straightforward manner via Fmoc SPPS and are later transformed into peptide azides which in turn form the thioester in presence of thiols. Experiments in solution proved successful with formation of thioester detected. When tried on a resin bound peptide however, the hydrazine peptides were completely inert to the treatment with no product formed. A third approach took advantage of the auxiliary inability to mediate ligation with benzylic thioesters to perform kinetically controlled ligation. The benzyl thioester MUC1, anchored to the His-Tag resin, was treated with a buffer containing MPAA, thiophenol or MESNa. MPAA and thiophenol did not produce the desired thioester. Treatment with a buffer containing MESNa, a full conversion was observed. However, the MESNa thioester proved not reactive enough to react with the auxiliary peptide, failing to produce the ligation product.




**Scheme 4.3**: Strategies for activation of masked thioesters..**A)** Using Dbz linker; **B)** Using peptide hydrazides; **C)** Using KCL.

Finally the applicability of the method was tested for ligating long peptides fragments by synthesizing a 80mer MUC1 VNTR domain peptide. The reaction was successful, demonstrating the auxiliary applicability to mediate ligation of long peptides. The ligation proved more challenging than the previous shorter fragment condensation, taking 3 days to complete and requiring addition of thioester peptide to be driven to completion. In cases such as these the advantage of solid support chemistry come into play, as the excess thioester, which would normally remain in solution and often complicates purification, was washed off the support contributing to high crude purity. As before, the auxiliary was removed in-situ to give the native MUC1 peptide (figure 4.7).

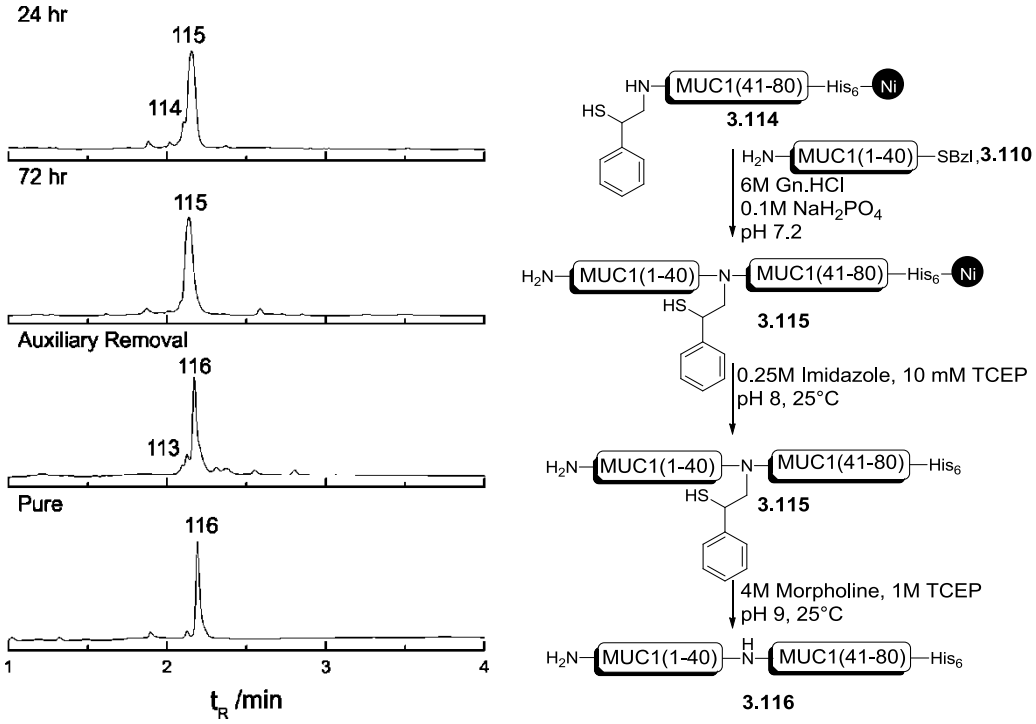



**Figure 4.7**: UPLC analysis of synthesis of 80mer MUC1 VNTR peptide **3.116**; 03-60% B in 6 min, 210 nm.

## 4.2 Outlook

This work has continued previous efforts to develop a new type of auxiliaries for extending native chemical ligation at cysteine free junctions [96]. Unlike the existing auxiliaries for NCL which are based on protecting group derivatives and are removed under acidic conditions, the new auxiliaries are based on an aliphatic flexible scaffold. Their removal chemistry was completely revamped to proceed under mild basic conditions. These changes were envisioned to allow these auxiliaries to overcome the problems encountered by the existing array of auxiliaries. The main features for optimization refer to expanding the scope of available ligation junctions and development of auxiliary removal conditions that are compatible with sensitive peptide moieties and modifications.

From observations made in this work several modifications to the system could be incorporated to improve the auxiliaries' usability. Results have shown that the β-branched substituted amine prevented ligation at sterically demanded junctions to proceed. Redesigning the auxiliary to produce a sterically less-demanding secondary amine will serve two purposes: more effective coupling of the auxiliary to the peptide and allow ligation of peptide fragments beyond the Ala-Gly and Gly-Ala junctions. A redesign of the scaffold could also prevent the undesired thiolactonisation, thus eliminating the auxiliaries' premature removal. The auxiliaries' removal chemistry, based on a mild basic treatment should theoretically prove compatible with common sensitive modifications of proteins. Here too, optimization of the EWG and removal conditions could improve the removal rate and selectiveness.

The results gained in the first part of the study have also led to the design of a "second-generation" auxiliary with highly improved properties in key factors, mainly: a more efficient attachment to the peptide and broader number of accessible ligation junctions. The auxiliary has also proven to be efficient in solid phase chemical ligation in a variety of peptides lengths, giving access to highly pure products. Further improvements and modification of the auxiliary such as derivatization of the phenyl ring with electron donating groups should improve the rate of ligation in sterically hindered junctions as well as allowing faster, more selective and more efficient removal of the auxiliary to achieve the native protein.

Despite difficulties in performing sequential SPCL, it is not an inherent problem but rather a question of utilizing the appropriate method. Other latent thioesters than the one tested could be used, such as thiocarboxylates [92], SEA linker [93] or system based on a N→S or O→S acyl shift [85, 88, 90, 137-139]. Alternatively, orthogonal thiol protecting groups like thiazolidine

or 2-(methylsulfonyl)ethyl carbonate could be utilized to perform sequential ligations in the C→N direction.

A possible application for auxiliary assisted SPCL could take advantage of the inherited purification characteristic of solid phase reactions. By performing the ligation in the N→C direction, a double purification process could be achieved without the need for HPLC purification. First, the crude His-tagged thioester could be anchored on the solid support. All truncation could be washed off the support leaving only the full length tagged peptide anchored (one must take care that no truncations in the His-Tag chain occurs which can be done for example by employing the "thioester with self-purification" method [140]). The crude auxiliary peptide could be directly ligated as all truncated fragments will lack the auxiliary moiety and will be washed off the support leaving the pure ligation product as a single species.

The initial demonstration of the auxiliary SPCL capabilities should open a door to synthesis of longer and more complex proteins of interest by sequential SPCL with high purity and yield, hopefully turning auxiliary-assisted NCL wide-spread and useful addition to the toolkit of scientist working at the edge of biology and chemistry.

# 5 Experimental Part

## 5.1 Instruments and Materials

**Chemicals and Instruments Specifications:** Commercially available compounds were used without further purification. Dry solvents were taken from a *Braun* solvent purification system *SPS 800*. Purification of compounds by flash chromatography was done on silica gel (0.060-0.2 mm, 60 Å) from *Acros Organic*s using technical grade solvents. TLC silica gel plates 60 $F_{254}$ from *Merck* were used for thin-layer chromatography. NMR-spectra were recorded on a *Bruker Avance II 500 MHz* spectrometer and referenced to the residual protonated solvent signal. HRMS-spectra were recorded on a *QstarXL* spectrometer from *Applied Biosystems*.

**SPPS:** Solid-phase chemistry was carried out in syringes, equipped with teflon filters, purchased from Torviq or using peptide synthesizer (CS Bio Co. CS336x). If not described differently, all reactions were carried out at room temperature. DMF was purchased in biotech grade. Peptide Commercial reagents were used without further purification. Resins, protected amino acids and activating agents were purchased from Novabiochem.

**UPLC-MS** measurements were performed on an Acquity UPLC system from Waters using a BEH130 C18 column (130 Å, 1.7 µm, 2.1 x 50 mm; heater set on 50 °C) with a binary mixture of A (0.1 % TFA, 1 % acetonitrile, 98.9 % $H_2O$) and B (0.1 % TFA, 1 % $H_2O$, 98.9 % acetonitrile) as a mobile phase (flow = 0.5 mL/min) in a linear gradient as described.

**HPLC-MS** measurements were performed on an 1100 Series system from Agilent using a Polaris C18 column (3 µm, 2 X 100 mm; heater set on 50 °C) from Varian with a binary mixture of A (0.1 % formic acid, 1 % acetonitrile, 98.9 % $H_2O$) and B (0.1 % formic acid, 1 % $H_2O$, 98.9 % acetonitrile) as a mobile phase (flow = 0.3 mL/min) in a linear gradient as described.

**Preparative HPLC** purifications were carried out on a 1100 Series system from Agilent using a Nucleodur C18 Gravity (5 µm, 250 mm) from Macherey-Nagel with a binary mixture of A (0.1 % TFA, 1 % acetonitrile, 98.9 % $H_2O$) and B (0.1 % formic acid, 1 % $H_2O$, 98.9 % acetonitrile) as a mobile phase (flow = 15 mL/min) in a linear gradient as described. For semi-preparative HPLC-purifications a Polaris C18-A (5 µm, *250* x 10.0 mm) from Varian was used with a binary mixture of A (0.1 % TFA, 1 % acetonitrile, 98.9 % $H_2O$) and B (0.1 % TFA, 1 % $H_2O$, 98.9 % acetonitrile) as a mobile phase (flow = 6.0 mL/min) in a linear gradient as described.

**Concentrations of peptide solutions** were determined by using the absorption at 278 nm measured with ND-1000 spectrophotometer from Peqlab Biotechnology GmbH. The extinction coefficients of the peptides were calculated as the sum of the extinction coefficients of the tyrosine and tryptophane residues:

$\varepsilon 278(Tyr) = 1400\ L/(mol \cdot cm)$

$\varepsilon 278(Trp) = 5600\ L/(mol \cdot cm)$

**Coupling Efficiency Determination**: Fmoc group is removed as described before. The combined piperidine washes (1 ml) are united. 20% piperidine in DMF (900 µl) in a UV cuvette is placed in a spectrophotometer and blank is measured at λ=301 nm. 100 µl of the Fmoc cleavage solution is added to the cuvette, mixed and the absorbance is measured at 301 nm. Coupling is calculated using the equation: n (mmol) = $\frac{A \times Dilution \times V_{piperidine\ solution}(ml)}{l_{cell}(cm) \times \varepsilon(cm^{-1}M^{-1})}$

# 5.2 General Methods

**Manual SPPS by Fmoc Strategy**: Amino acid, HCTU and HOBt (4 eq. each) and NMM (8 eq.) were dissolved in DMF (final concentration 0.4 M) and transferred to the resin. After 30 min, the resin was washed with DMF (X5) and unreacted sites were capped by treating the resin with 1 ml solution of 5 % $Ac_2O$ and 6 % 2,6-lutidine in DMF (1:2:1 min). Resin was washed with DMF (X5). Fmoc Group was cleaved by treating the resin with 1 ml solution of 20% piperidine in DMF (3:7:3 min).

**Synthesis of Thioesters by Boc Strategy**: MBHA resin (0.59 mmol/g) was washed with DMF (X3) and treated with 10% DIEA/DMF for 10 min and additional wash with DMF (X3). Fmoc-Gly-OH (4 eq) was coupled on a HMBA resin using HBTU and HOBt (4 eq. each) and NMM (8 eq.) in DMF (final concentration 0.4 M). Fmoc was removed by treating the resin with 1 ml solution of 20% piperidine in DMF (3:7:3 min) and loading was determined. 3-(tritylthio) propanoic acid (4 eq) was pre-activated with HBTU (4eq) and DIEA (8 eq) in DMF (final concentration 0.4 M) for 5 min and added to the resin. After 1 hr the resin was washed with DMF (X3). The trityl protecting group was removed using a mixture of TFA:TIS:water (90:5:5 v/v). The *in situ* neutralization protocol developed by Kent and coworkers was applied for peptide assembly [141].

*Boc deprotection*: The resin was washed with DCM (X5) and treated with 1 ml solution of 5% m-cresol in TFA (1:2:1 min). Resin is washed with DCM(X5) then DMF (X5).

*Cleavage*: A mixture of TFMSA:TFA:thioanisole (2:8:1, v/v) was added to the resin. After 1.5 h, the resin was washed with TFA (4 x 4 mL).

*Work-up*: The combined solutions were concentrated in vacuum. The residue was dissolved in water, purified by preparative HPLC and analyzed by HPLC-MS.

**Automated SPPS**: The reaction vessel containing the preloaded resin was placed in an automated peptide synthesizer.

*Coupling*: Amino acid, HCTU, OxymaPure (4 eq. each) and NMM (8 eq.) are mixed together in DMF (final concentration 0.4 M) and transferred to the resin.
Coupling times: Residues 1-10 1X 30 min; Residues 11-24 2X30 min; Residues 25-End 3X30 min.
Between couplings the resin was washed with DMF (3X500 µl).

*Capping*: A mixture of 5 % $Ac_2O$ and 6 % 2,6-lutidine in DMF (400 µl) was added to the resin for 5 min.

*Fmoc deprotection*: Resin was treated with 500 µl 20% piperidine in DMF.
Deprotection times: Residues 1-10 5:5 min; Residues 11-End 5:8 min.

After last cycle the resin is capped and washed with DMF (X3) and DCM (X3) and dried under vacuum.

**Thioester Synthesis using Safety-Catch Linker Strategy**:

*Loading*: 4-sulfamylbutyryl resin (5 eq.) swelled in chloroform. Amino Acid, PyBOP (1 eq.) and DIEA (3 eq.) dissolved in $CHCl_3$ (final concentration 0.25-0.3 M) and added to resin. The resin is shaken at -20 °C for 2 hr then washed with $CHCl_3$ and swelled in DMF.

*Loading determination*: Performed by monitoring loading efficiency as described below.

Rest of synthesis is continued on the automatic peptide synthesizer as described above.

*Changing the N-termini protecting group:* Fmoc is removed as described in the general protocol and coupling efficiency is checked as described below. $Boc_2O$ and DIEA (50 eq. each ) in enough DMF to cover the resin is added and shaken at rt for 30 min then washed with DMF and DCM.

*Activation of sulfonamide linker*: Resin swelled in dry THF. Trimethylsilyldiazomethane (2 M in hexane) (50 eq. from original loading) is diluted with dry THF 1:1v/v and added to resin at rt for 2 hr. Resin washed with THF and DCM.

*Thiolysis and release from resin*: Resin swelled in DMF. Benzyl mercaptan (50 eq. from original loading) and sodium thiophenolate (1 eq. from original loading) in dry DMF are added to resin. Resin is shaken at rt 16 hr. The solution is transferred into a falcon tube and resin is washed thoroughly with DMF (X5) and DCM (X5). The combined washes are added to the falcon tube and solution is evaporated to dryness under vacuum.

*Cleavage of protecting groups and isolation of peptide*: Performed as described below under the general procedure for Fmoc SPPS cleavage and purification of peptides.

**Thioester synthesis using N-Acylurea Strategy**:

*Preparing the Dbz modified resin*: Rink amide resin (0.19 mmol/g, 0.05 mmol scale) was swelled in
DMF. Fmoc-Gly-OH (0.2 mmol, 4 eq) and Fmoc-3,4-diaminobenzoic acid (0.2 mmol, 4 eq) were double coupled for 1 hr. Resin is capped as described above.

*Alloc protection*: The resin is swelled in DCM allylchloroformate (5 mmol, 100 eq) and DIEA (0.05 mmol, 1eq) in dry DCM are added to the resin for 14 hr. Resin is washed with DCM and DMF and capped.

*Loading of first amino acid*: Fmoc-His(Trt)-OH (0.2 mmol, 4 eq), HATU (0.2 mmol, 4eq) and NMM (0.4 mmol, 8 eq) are dissolved in DMF and added to resin for 2 X 1 hr.

Rest of synthesis is continued on the automatic peptide synthesizer as described above.

*Changing the N-termini protecting group:* Fmoc is removed as described in the general protocol and
coupling efficiency is checked as described below. $Boc_2O$ and DIEA (50 eq. each ) in enough DMF to cover the resin is added and shaken at rt for 30 min then washed with DMF and DCM.

*Alloc Deprotection*: Resin is swelled in dry DCM and $Pd(PPh_3)_4$ (0.15 mmol, 3 eq) in 92.5:5:2.5 v/v
DCM:AcOH:NMM is added to resin and shaken 2 hr at rt. Resin is washed with DCM, DMF, 0.5% v/v DIEA/DMF, 0.5% w/v diethyldithiocarbonate/DMF, DMF and DCM.

*Cyclization*: *p*-nitrophenylchloroformate (0.25 mmol, 5 eq) in dry DCM was added to resin and the resin was shaken for 1 hr at rt. Resin was washed with DCM and swelled in DMF and treated with 0.5 M DIEA in DMF for 2 x 15 min.

**Reductive Amination on Solid Support**

*Method A*: The auxiliary (20 eq.) is dissolved in a mixture of 1:1:1 NMP:MeOH:TMOF +5% AcOH (final concentration 0.5 M) and added to resin. After 90 min sodium cyanoborohydride (20 eq.) in 1:1 NMP:MeOH (final concentration 0.5 M) is added and the reaction is shaken at rt for 16 hr. Procedure is repeated if needed.

*Method B*: The auxiliary (200 eq.) is dissolved in a mixture of 1:1:1 NMP:MeOH:TMOF +5% AcOH (final concentration 0.5 M) and added to resin. After 90 min sodium cyanoborohydride (200 eq.) in 1:1 NMP:MeOH (final concentration 0.5 M) is added and the reaction is shaken at rt for 16 hr. Procedure is repeated if needed.

*Method C*: The auxiliary (20 eq.) and sodium cyanoborohydride (20 eq.) are dissolved in a mixture of 3:1 NMP:iPrOH +5% AcOH (final concentration 0.4 M) and added to resin for 4-16 hr.

**Cleavage from Support and Protecting Groups Removal**

*Fmoc Strategy:* The dry peptidyl resins or the dry thiolysis residue (250 mg) were treated with 2 ml 95:2.5:2.5 v/v TFA:$H_2O$:TIS solution for 2 hr. After 2 hr the mixture was transferred to a tube. The resin was washed with the cleavage mixture (0.5 ml X 5) and combined washes were united, concentrated under an argon flow to ~2 ml and added dropwise to cold diethylether (20 ml). The mixture was centrifuged for 15 min at 0 °C. The ether was decanted and the precipitate was dissolved in 25 % $CH_3CN$ in $H_2O$ + 0.1% TFA and lyophilised to dryness.

*Boc Strategy:* The dry peptidyl resin (250 mg) is placed in a RBF with a stirring bar. A solution of 750 µl 2:1 v/v thioanisole: EDT. RBF is cooled in an ice bath and 5 ml TFA is added for 10 min followed by drop-wise addition of 500 µl TFMSA. After 2 hr the mixture was collected and resin washed with TFA (5X 0.5 ml). Combined washes were concentrated under an argon flow to ~ 2

ml and added drop-wise to cold ether (20 ml) and centrifuged for 15 min at 0 °C. The ether was decanted and the precipitate was dissolved in 25 % $CH_3CN$ in $H_2O$ + 0.1% TFA and lyophilised to dryness.

**tert-Butyl Deprotection from Peptides 3.034-3.041**

*Method A*: The dry peptidyl resin (250 mg) is placed in a round bottom flask with a stirring bar. A solution of 750 µl 2:1 v/v thioanisole: EDT. RBF is cooled in an ice bath and 5 ml TFA is added for 10 min followed by drop-wise addition of 500 µl TFMSA. After 2 hr the mixture was collected and resin washed with TFA (5X 0.5 ml). Combined washes were concentrated under an argon flow to ~ 2 ml and added drop-wise to cold ether (20 ml) and centrifuged for 15 min at 0 °C. The ether was decanted and the precipitate was dissolved in 25 % $CH_3CN$ in $H_2O$ + 0.1% TFA and lyophilised to dryness.

*Method B*: The crude lyophilized peptide is dissolved in cold TFA (concentration of 0.05-0.1 M) and placed in an ice-bath. Mercury acetate (2 eq.) is added and mixture is left at 0 °C for 30 min. Dithiothreitol (4 eq.) is added and mixture is left at 0 °C for another 10 min. The mixture is centrifuged and the supernatant is collected. The mercuric sulphide salt precipitate is washed with TFA, centrifuged and supernatant is collected again. The procedure is repeated 3 times and the united TFA washes are concentrated, diluted to 900 µl with water and purified by preparative HPLC.

**Native Chemical Ligation**

*Method A*: thioester and auxiliary peptide in $H_2O$ +0.1% TFA solution are mixed together in a ratio 1.2-1.4:1 respectively and lyophilized. The solid mixture is dissolved in a buffer containing 0.1 M $NaH_2PO_4$, 20 mM TCEP, 3%v/v PhSH, pH 7.5. pH is adjusted to 7.5 with 0.1 M NaOH (final concentration 5 mM) and reaction is shaken at 25 °C. The progress of reaction was monitored using UPLC-MS. Ligation product was isolated by semi-preparative HPLC.

*Method B*: thioester and auxiliary peptide in $H_2O$ +0.1% TFA solution are mixed together in a ratio of 1.2:1 respectively and lyophilized. The solid mixture is dissolved in a buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP, 2% v/v PhSH, and BzlSH pH 7.5. pH is adjusted to 7.5 with 0.1 M NaOH (final concentration 5 mM) and reaction is incubated at 25 °C. The progress of reaction was monitored using UPLC-MS. Ligation product was isolated using semi-preparative HPLC.

**Auxiliary Removal**

*Method A*: The pure auxiliary modified ligation product was dissolved in an aqueous buffer of 140 mM morpholine and 20 mM TCEP pH 8.5 to a concentration of 2 mM and incubated at 60 °C for 6 hr. The reaction is quenched by diluting it with $H_2O$ + 0.1% TFA and directly purified by semi-preparative HPLC.

*Method B*: The pure auxiliary modified ligation product was dissolved in 50 µl aqueous buffer containing 400 mM morpholine and 100 mM TCEP pH 8.5 to a concentration of 2 mM and incubated at 50 °C for 6 hr. The reaction is quenched by diluting it with $H_2O$ + 0.1% TFA and directly purified by semi-preparative HPLC.

*Method C*: To the crude auxiliary modified ligation product in 100 µl aqueous buffer of 250 mM imidazole and 10 mM TCEP is added 100 µl aqueous buffer containing 400 mM morpholine and 100 mM TCEP pH 8.5 (final concentration 0.05 mM) and incubated at 25 °C for 24 hr. Reaction is quenched by diluting with $H_2O$ + 0.1% TFA and directly purified by semi-preparative HPLC.

**Auxiliary-Assisted Solid Phase Chemical Ligation**

*Loading Ni-NTA resin with His-Tag peptides*: The Ni-NTA resin (50% suspension in 30% EtOH:$H_2O$, 1.2 eq., 70 mg/ml) was swelled in a buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$ pH 7.2 for 30 min. His-Tag peptide (1 eq.) dissolved in same buffer was added to resin for 60 min and resin is washed with the buffer (X5).

*C→N directed SPCL*: To the Ni-NTA resin bound auxiliary peptide (1 eq.) was added the thioester peptide fragment (1.2 eq.) dissolved in a buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP, 10 mM MPAA pH 7.2 to a concentration of 5 mM and reaction was shaken at 25 °C for 16 hr. The process was repeated if needed using 0.5 eq. of thioester peptide.

*N→C directed SPCL*: To the Ni-NTA resin bound thioester peptide (1 eq.) was added the auxiliary peptide (1.2 eq.) dissolved in a buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP, 10 mM MPAA pH 7.2 to a concentration of 5 mM and reaction was shaken at 25 °C for 16 hr. The process is repeated if needed using 0.5 eq. of auxiliary peptide.

*Elution of His-Tag peptides from Ni-NTA resin*: The resin was thoroughly washed with buffer containing 6 M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP pH 7.2, water, 0.1% TFA in

water, and water. Elution buffer containing 250 mM imidazole and 10 mM TCEP pH 8.0 (40 µl) was added for 30 min. The elution buffer was collected and process is repeated again. The resin was then washed with the elution buffer (2 X 10 µl) and all washes are combined.

**Transthioesterification on Ni-NTA Resin**: the resin bound peptide benzyl thioester was treated with a buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 40 mM thiol, pH 7.0 and incubated at 25 °C for 16 hr.

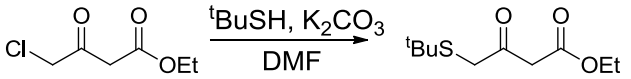
**Converting Peptide Hydrazides to Peptide Thioesters**:

*Method A*: The resin bound peptide hydrazide was treated with a buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 30 mM $NaNO_2$ pH 3.0 at -10 °C for 20 min, washed and treated with buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM thiol pH 7.2 at 25 °C for 30 min.

*Method B*: Peptide hydrazide was dissolved in buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 30 mM $NaNO_2$ pH 3.0 at -10 °C for 20 min. To the mixture a buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM thiol pH 7.2 is added at 25 °C for 30 min.

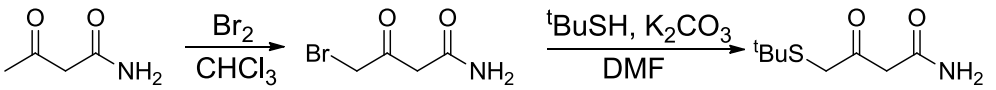
# 5.3 Experimental Data for NCL Assisted by Base-Labile Auxiliary

## 5.3.1 Synthesis of Auxiliary Precursors




**3.020** **3.021** 85%

**Ethyl 4-tertbutylsulfonyl-3-oxobutanoate (3.021)**: $K_2CO_3$ (1 gr, 7.3 mmol, 1.2 eq) was suspended in 12 ml DMF. Tert-butylthiol (820 µl, 7.3 mmol,, 1.2 eq) is added and the mixture is stirred at rt for 30 min. Ethyl 4-chloro-3-oxobutanoate **3.020** (1 gr, 900 µl, 6.1 mmol, 1.21 gr/ml, 1 eq), is added and mixture is further stirred at rt for 3 hr. Ethyl acetate (15 ml) and water (15 ml) are added to the reaction. Phases are separated and organic phase is washed with water (3 X 10 ml) and brine (5 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure to yield pure ethyl 4-tertbutylsulfonyl-3-oxobutanoate **3.021** with quantitative yield (1.3 gr, 6 mmol). $^1$H NMR ($CDCl_3$, 300 MHz): δ= 4.18 (q, 2H, J= 7.2 Hz , $OC\underline{H_2}CH_3$), 3.68 (s, 2H, $^t$BuS$C\underline{H_2}$C(O)), 3.43 (s, 2H, C(O)$C\underline{H_2}$C(O)), 1.30 (s, 9H, $C(CH_3)_3$), 1.26 (t, 3H, J= 7.1 Hz , $OCH_2C\underline{H_3}$) $^{13}$C NMR ($CDCl_3$, 75 MHz): δ= 200.5, 167.3, 62.0, 46.5, 43.6, 39.7, 30.7, 14.0 0. ESI-MS ($CHCl_3$: MeOH) m/z: Calculated for $[C_{10}H_{19}O_3S]^+$: 219.1049 Da $[C_{10}H_{18}O_3SNa]^+$: 241.0869 Da. Observed: 219.1051 $[M+H^+]$, 241.0871 $[M+Na^+]$.

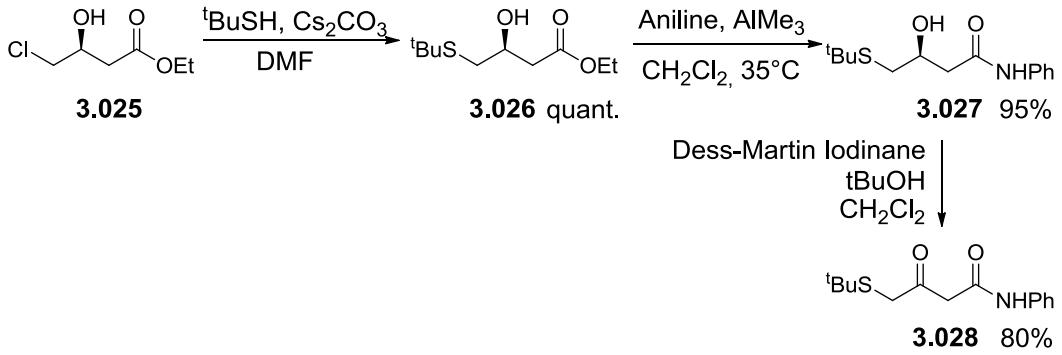



**3.022** **3.023** **3.024** 63%

**4-bromoacetoacetamide (3.023)**: Acetoacetamide **3.022** (500 mg, 4.95 mmol, 1 eq) is suspended in 1.5 ml $CHCl_3$ (C=3.3 M) and cooled in an ice bath to 0 °C. Bromine (280 µl, 5.45 mmol, 1.1 eq) in 1 ml $CHCl_3$ is added drop-wise and mixture is further stirred at 0 °C for 2 hr. Ice is added to the reaction and product is extracted with chloroform (5 X 3 ml). Combined organic phases are dried over $Na_2SO_4$, filtered and concentrated under reduced pressure to yield 4-bromoacetoacetamide **3.023** as dark brown oil that was uses without further purification.

**4-tertbutylsulfonylacetoacetamide (3.024)**: $K_2CO_3$ (825 mg, 5.95 mmol, 1.2 eq) was suspended in 5 ml DMF. Tert-butylthiol (670 µl, 5.95 mmol, 1.2 eq) is added and the mixture is stirred at rt for 10 min. **3.023** (4.95 mmol, 1 eq**)** is dissolved in 5 ml DMF and added to the reaction mixture.

Reaction is allowed to stir at rt for 2 hr. Ethyl acetate (15 ml) and water (15 ml) are added to the reaction. Phases are separated and organic phase is washed with water (3 X 10 ml) and brine (5 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. Column purification yielded pure 4-tertbutylsulfonylacetoacetamide **3.024** as a red oil (234 mg, 1.24 mmol, 25%). $^1$H NMR (MeOD-D4, 300 MHz): δ= 3.01 (s, 2H, C(O)C$\underline{H_2}$C(O)), 2.88 (2, 2H, -SC$\underline{H_2}$C(O)), 1.35 (s, 4.5H, C(C$\underline{H_3}$)$_3$) $^{13}$C NMR (MeOD-D4, 75 MHz): δ= 163.5, 35.6, 30.2, 29.7 ESI-MS (MeOH) m/z: Calculated for $[C_8H_{16}NO_2S]^+$: 190.0896 Da $[C_8H_{15}NO_2SNa]^+$: 212.0716 Da $[C_8H_{15}NO_2SK]^+$: 228.0455 Da Observed: 190.0898 $[M+H^+]$, 212.0718 $[M+Na^+]$, 228.0458 $[M+K^+]$.

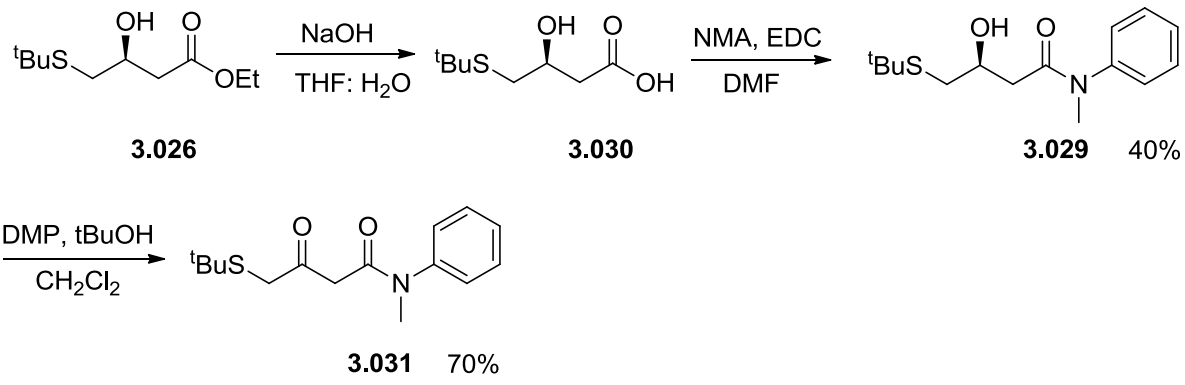



**(S)-Ethyl 4-tertbutylsulfonyl-3-hydroxybutanoate (3.026)**: $Cs_2CO_3$ (2.4 gr, 7.5 mmol, 2.5 eq.) was suspended in 6 ml DMF. Tert-butylthiol (760 µl, 7.5 mmol, 2.5 eq.) is added and the mixture is stirred at rt for 30 min. Ethyl 4-chloro-3-hydroxybutanoate **3.025** (420 µl, 3 mmol, 1 eq.) is added and mixture is further stirred at rt for 3 hr Ethyl acetate (10 ml) and water (10 ml) are added to the reaction. Phases are separated and organic phase is washed with water (3 X 10 ml) and brine (5 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure to yield pure ethyl 4-tertbutylsulfonyl-3-hydroxybutanoate **3.026** as colorless oil with quantitative yield (635 mg, 2.9 mmol). $^1$H NMR ($CDCl_3$, 500 MHz): δ= 4.19 (q, 2H, J= 7.2 Hz , OC$\underline{H_2}$CH$_3$), 3.17 (d, 1H, J= 3.8 Hz, CH$_2$C$\underline{H}$OH), 2.63 (m, 4H, C$\underline{H_2}$CHOHC$\underline{H_2}$)), 1.35 (s, 9H, C(CH$_3$)$_3$), 1.29 (t, 3H, J= 7.1 Hz , OCH$_2$C$\underline{H_3}$) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 172.2, 67.5, 60.8, 42.6, 40.4, 35.0, 31.0, 14.2 ESI-MS (MeOH) m/z: Calculated for $[C_{10}H_{21}O_3S]^+$: 221.1206 Da $[C_{10}H_{21}O_3S\ Na]^+$: 243.1025 Da $[C_{14}H_{21}NO_2S\ K]^+$: 259.0975 Da Observed: 221.1209 $[M+H^+]$, 243.1029 $[M+Na^+]$, 259.0979 $[M+K^+]$.

**(S)-4-tertbutylsulfonyl-3-hydroxy-N-phenylbutanamide (3.027)**: Aniline (228 µl, 2.5 mmol, 2.5 eq.) is dissolved in 7 ml $CH_2Cl_2$ and cooled to 0 °C. 2 M Trimethylaluminum in heptane (1.25 ml, 2.5 mmol, 2.5 eq.) is added and mixture is stirred at 0 °C for 1 hr. **3.026** (220 mg, 1 mmol, 1 eq.) is dissolved in 3 ml $CH_2Cl_2$ and added drop-wise to the mixture and the reaction is refluxed for 2 hr. Reaction is cooled to 0 °C, diluted with 10 ml $CH_2Cl_2$, cold 3 M aqueous HCl (20 ml) is

added drop-wise and stirred for 1 hr. Layers are separated and aqueous layer is washed with $CH_2Cl_2$ (3 X 10 ml). Combined organic layers are dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. The crude is then purified using column chromatography (3:7 EtOAc: cHex) to yield **3.027** as a white solid. (250 mg, 0.95 mmol, 95% yield) $^1$H NMR ($CDCl_3$, 500 MHz): δ= 8.10 (s(br), 1H, NHPh) 7.53 (d, 2H, J= 7.5 Hz , 2XArH), 7.33 (t, 2H, J= 7.5 Hz , 2XArH), 7.14 (t, 1H, J= 7.5 Hz , ArH), 4.15 (m, 1H, CHOH), 3.70 (s(br), 1H, OH), 2.70 (m, 4H, $CH_2CHOHCH_2$), 1.36 (s, 9H, $C(CH_3)_3$) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 169.5, 137.7, 129.0, 124.4, 120.0, 67.6, 43.1, 43.0, 35.6, 31.1 ESI-MS ($CHCl_3$: MeOH) m/z: Calculated for $[C_{14}H_{22}NO_2S]^+$: 268.1366 Da $[C_{14}H_{21}NO_2S\ Na]^+$: 290.1185 Da $[C_{14}H_{21}NO_2S\ K]^+$: 306.0925 Da Observed: 268.1368 $[M+H^+]$, 290.1187 $[M+Na^+]$, 306.0928 $[M+K^+]$.

**(S)-4-tertbutylsulfonyl-3-oxo-N-phenylbutanamide (3.028)**: Dess-Martin periodinane 15% weight solution in $CH_2Cl_2$ (3.2 ml, 1.7 mmol, 1.05 eq) is dissolved in 12 ml $CH_2Cl_2$. Tert-butanol (164 µl, 1.72 mmol, 1.06 eq) is added and mixture is stirred at rt for 20 min. **3.027** (430 mg, 1.62 mmol, 1 eq.) is dissolved in 3 ml $CH_2Cl_2$ and added to the mixture and reaction is allowed to stir at rt for 20 min. Reaction is diluted with ether (15 ml) and poured into a solution of 0.3 N sodium thiosulfate pentahydrate in 1 N sodium bicarbonate (60 ml) and stirred for 20 min. Layers are separated and organic phase is washed with 1N $NaHCO_3$ (15 ml), water (15 ml) and brine (10 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. The crude is purified using column chromatography (1:1 EtOAc: cHex) to yield the desired product **3.028** as brown oil (345 mg, 1.30 mmol, 80% yield). $^1$H NMR ($CDCl_3$, 500 MHz): δ= 8.93 (s(br), 1H, NHPh) 7.55 (d, 2H, J= 7.5 Hz , 2XArH), 7.33 (t, 2H, $J_{1,2}$= 7.5 Hz, $J_{2,3}$= 8.3 Hz, 2XArH), 7.14 (t, 1H, J= 7.4 Hz, ArH), 3.81 (s, 2H, C(O)$CH_2$C(O)), 3.48 (s, 2H, -S$CH_2$C(O)), 1.35 (s, 9H, $C(CH_3)_3$) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 204.2, 163.4, 137.5, 129.0, 124.6, 120.1, 47.3, 43.9, 40.5, 30.7 ESI-MS ($CHCl_3$: MeOH) m/z: Calculated for $[C_{14}H_{20}NO_2S]^+$: 266.1209 Da $[C_{14}H_{21}NO_2S\ Na]^+$: 288.1029 Da $[C_{14}H_{21}NO_2S\ K]^+$: 304.0768 Da Observed: 266.1212 $[M+H^+]$, 288.1032 $[M+Na^+]$, 304.0071 $[M+K^+]$.



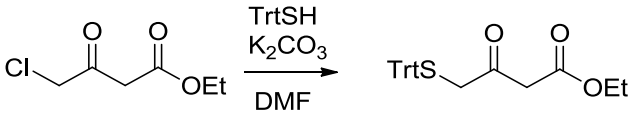

**3.026** **3.030** **3.029** 40%

**3.031** 70%

**(S) - 4-tertbutylsulfonyl-3-hydroxybutanoic acid (3.030)**: **3.026** (700 mg, 3.20 mmol) is dissolved in 9 ml THF and 5 ml aqueous solution of 2 N NaOH is added. Reaction is allowed to stir at rt for 3 hr then quenched with cold 2 N aqueous HCl solution to pH 2. Product is extracted with EtOAc (3 X 10 ml) and combined organic phases are washed with brine (5 ml), dried over $Na_2SO_4$ and concentrated to yield desired product **3.030** that was used without further purification.

**(S) - 4-tertbutylsulfonyl-3-hydroxy-N-methyl-N-phenylbutamide (3.029)**: **3.030** (340 mg, 1.5 mmol, 1 eq.) and N-methylaniline (196 µl, 1.8 mmol, 1.2 eq) are dissolved in 3 ml DMF and EDC.HCl (345 mg, 1.8 mmol, 1.2 eq) is added. Reaction is stirred at rt O.N. Ethyl acetate (5 ml) is added and washed with water (3 X 5 ml). Organic phase is dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. Purification with column chromatography (1:1 EtOAc: cHex) gave **3.029** (168 mg, 0.60 mmol, 40% yield over 3 steps). $^1$H NMR ($CDCl_3$, 500 MHz): δ= 7.42 (m, 3H, 3XAr$\underline{H}$) 7.22 (d, 2H, J= , 2XAr$\underline{H}$), 4.07 (m, 1H, C$\underline{H}$OH), 3.30 (s, 3H, NC$\underline{H_3}$Ph), 2.40 (m, 4H, C$\underline{H_2}$COHC$\underline{H_2}$), 1.28 (s, 9H, C(C$\underline{H_3}$)$_3$) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 171.7, 142.8, 129.5, 127.7, 126.8, 68.0, 41.9, 38.9, 36.7, 34.1, 30.5 ESI-MS ($CHCl_3$: MeOH) m/z: Calculated for $[C_{15}H_{24}NO_2S]^+$: 282.1522 Da $[C_{15}H_{23}NO_2SNa]^+$: 304.1342 Da $[C_{14}H_{21}NO_2SK]^+$: 320.1081 Da Observed: 282.1526 $[M+H^+]$, 304.1345 $[M+Na^+]$, 320.1086 $[M+K^+]$.

**4-tertbutylsulfonyl-3-oxo-N-methyl-N-phenylbutamide (3.031)**: Dess-Martin periodinane 15%weight solution in $CH_2Cl_2$ (3.1 ml, 1.60 mmol, 1.05 eq) is dissolved in 10 ml $CH_2Cl_2$. Tert-butanol (157 µl, 1.62 mmol, 1.06 eq) is added and mixture is stirred at rt for 20 min. **3.029** (430 mg, 1.53 mmol, 1 eq.) is dissolved in 4 ml $CH_2Cl_2$ and added to the mixture. Reaction is allowed to stir at rt for 20 min. Reaction is diluted with ether (15 ml) and poured into a solution of 0.3 N sodium thiosulfate pentahydrate in 1 N sodium bicarbonate (60 ml) and stirred for 20 min. Layers are separated and organic phase is washed with 1N $NaHCO_3$ (10 ml), water (10 ml) and brine (5 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. The crude is purified using column chromatography (1:1 EtOAc: cHex) to yield **3.031** as brown oil. (300 mg, 1.1 mmol, 70% yield). $^1$H NMR ($CDCl_3$, 500 MHz): δ= 7.42 (m, 3H, 3XAr$\underline{H}$), 7.25 (m, 2H, 2XAr$\underline{H}$), 3.52 (s, 2H, C(O)C$\underline{H_2}$C(O)), 3.35 (s, 2H, -SC$\underline{H_2}$C(O)), 3.51 (s, 3H, NC$\underline{H_3}$Ph), 1.24 (s, 9H, C(C$\underline{H_3}$)$_3$) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 202.0, 166.9, 143.5, 130.0, 128.3, 127.3, 46.5, 43.4, 39.8, 37.3, 30.7. ESI-MS ( $CHCl_3$: MeOH) m/z: Calculated for $[C_{15}H_{22}NO_2S]^+$: 280.1366 Da $[C_{14}H_{21}NO_2SNa]^+$: 302.1185 Da $[C_{14}H_{21}NO_2S\ K]^+$: 318.0925 Da Observed: 268.1369 $[M+H^+]$, 302.1189 $[M+Na^+]$, 318.0929 $[M+K^+]$.




**3.020** **3.070**

**Ethyl 3-oxo-4-(tritylthio)butanoate (3.070):** $K_2CO_3$ (423 mg, 3.04 mmol, 1 eq) was suspended in 6ml DMF. Triphenylmethylthiol (756 mg, 2.74, 0.9 eq) is added and the mixture is stirred at rt for 10 min. Ethyl -chloro-3-oxobutanoate **3.20** (410 µl, 3.04 mmol, 1 eq) is added and mixture is further stirred at rt for 3 hr. Ethyl acetate (15 ml) and water (15 ml) are added to the reaction. Phases are separated and organic phase is washed with water (3 X 10 ml) and brine (5 ml), dried over $Na_2SO_4$, filtered and concentrated under reduced pressure. Crude is purified by column (3:17 EtOAc/cHex) to yield pure ethyl 4-tertbutylsulfonyl-3-oxobutanoate **3.070** (984 mg ,2.43 mmol, 80%). $^1$H NMR ($CDCl_3$, 400 MHz): δ= 7.45-7.23 (m, 15H, 15X Ar<u>H</u>) 4.13 (q, 2H, J= 7.2 Hz , O<u>C$H_2$</u>$CH_3$), 3.26 (s, 2H, $^t$BuSC<u>$H_2$</u>C(O)), 3.21 (s, 2H, C(O)C<u>$H_2$</u>C(O)), 1.24 (t, 3H, J= 7.1 Hz , $OCH_2$C<u>$H_3$</u>) $^{13}$C NMR ($CDCl_3$, 125 MHz): δ= 198.5, 144.0, 129.5, 128.2, 127.0, 61.4, 47.3, 42.7, 14.0 ESI-MS ($CHCl_3$: MeOH) m/z: Calculated for $[C_{25}H_{24}NO_3S\ Na]^+$: 427.1338 Da $[C_{14}H_{21}NO_2S\ K]^+$: 443.1078 Da Observed: 427.1349 [M+$Na^+$], 443.1088 [M+$K^+$].

## 5.3.2 Synthesis of Model Peptides

**Synthesis of peptides GRAEYSGLG (3.32) and ARAEYSGLG (3.33)**: The peptides were manually synthesized as described in the general protocol on a Rink Amide resin (0.19 mmol/g) in a 100 µmol scale.

**3.32** H-GRAEYSGLG-$NH_2$: ESI-MS (positive mode): m/z = 909.0 $[M+H]^{+1}$, 455.0 $[M+2H]^{+2}$, Calculated Mass: 909.0 Da $[M+H]^{+1}$, 455.0 $[M+2H]^{+2}$.
ESI-MS (positive mode): m/z = 909.0 ($C_{38}H_{62}N_{13}O_{13}$ $[M+H]^{+1}$, calcd.: 737.9 ), 455.0 ($C_{38}H_{63}N_{13}O_{13}$ $[M+2H]^{+2}$, calcd.: 455.0).

**3.33** H-ARAEYSGLG- $NH_2$: Observed Mass: 923.0 g/mol $[M+H]^{+1}$, 462.0 $[M+2H]^{+2}$, Calculated Mass: 923.0 Da $[M+H]^{+1}$, 462.0 $[M+2H]^{+2}$.
ESI-MS (positive mode): m/z = 923.0 ($C_{39}H_{64}N_{13}O_{13}$ $[M+H]^{+1}$, calcd.: 923.0 ), 462.0 ($C_{39}H_{65}N_{13}O_{13}$ $[M+2H]^{+2}$, calcd.: 462.0).

**Synthesis of thioester peptides LYRAG-MPA-G (3.52) and LYRAA-MPA-G (3.53) using Boc SPPS**: MBHA resin (149.2 mg, 0.67 mmol/gr, 0.1 mmol) is swelled for 30 min in DMF, treated with 10% v/v solution of DIEA in DMF (2 X 5 min, 1 ml) and washed with DMF (X5). Fmoc-Gly-OH (70 mg, 0.4 mmol, 4 eq.), HCTU (165 mg, 0.4 mmol, 4 eq), HOBt (54 mg, 0.4 mmol, 4 eq) and NMM (88 µl, 0.8 mmol, 8 eq.) in 1 ml DMF is added to resin and coupled for 1 hr. Fmoc deprotection was performed and loading was measured to be quantitative. Resin is

washed with DMF (X5) and Trt-MPA-OH (139.4 mg, 0.4 mmol, 4 eq.), HCTU (165 mg, 0.4 mmol, 4 eq), HOBt (54 mg, 0.4 mmol, 4 eq) and NMM (88 µl, 0.8 mmol, 8 eq.) in 1 ml DMF is added to resin and coupled for 2 hr. Resin is washed with DMF (X5), capped, washed with DMF (X5) and DCM (X10), dried and weighted. Divided into 2 reaction vessels (98 mg, 50 µmol) and swelled in DCM. Trt group is deprotected by treatment with a mixture of 90:5:5v/v TFA:TES:$H_2O$ (2X5 min, 1 ml). Resin washed with DCM (X10) and DMF (X5). Rest of the sequence was synthesized as described in the general protocol.

**3.52** H-LYRAG-MPA-G-$NH_2$: 21.7 mg, 722.8 g/mol, 30 µmol, 60%. $C_{31}H_{51}N_{11}O_7S$, 722.88 g/mol. UPLC: $t_R$ = 1.10 min (03% B – 60% B in 2 min, ESI-MS (positive mode): m/z = 723.8 ($C_{31}H_{52}N_{11}O_7S$ $[M+H]^{+1}$, calcd.: 723.9 ), 362.4 ($C_{31}H_{53}N_{11}O_7S$ $[M+2H]^{+2}$, calcd.: 362.4).

**3.53** H-LYRAA-MPA-G-$NH_2$: 22.0 mg, 736.9 g/mol, 30 µmol, 60%.

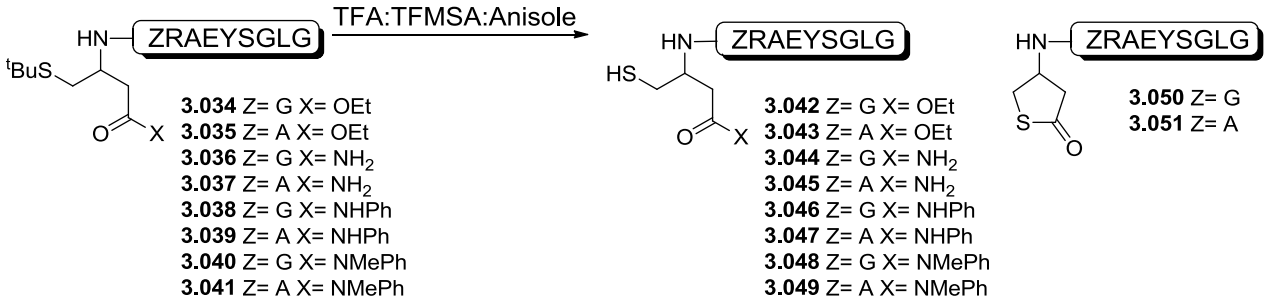
ESI-MS (positive mode): m/z = 737.9 ($C_{32}H_{54}N_{11}O_7S$ $[M+H]^{+1}$, calcd.: 737.9 ), 369.4 ($C_{32}H_{55}N_{11}O_7S$ $[M+2H]^{+2}$, calcd.: 369.4).

## 5.3.3 Reductive Amination on Solid Support

Auxiliary precursors **3.021**, **3.024**, **3.028** and **3.031** were attached to peptides **3.032** and **3.033** following the general reductive amination protocol method A.

## 5.3.4 tert-Butyl Protecting Group Removal

**Removal of tert-Butyl using method A: TFMSA:TFA:anisole**



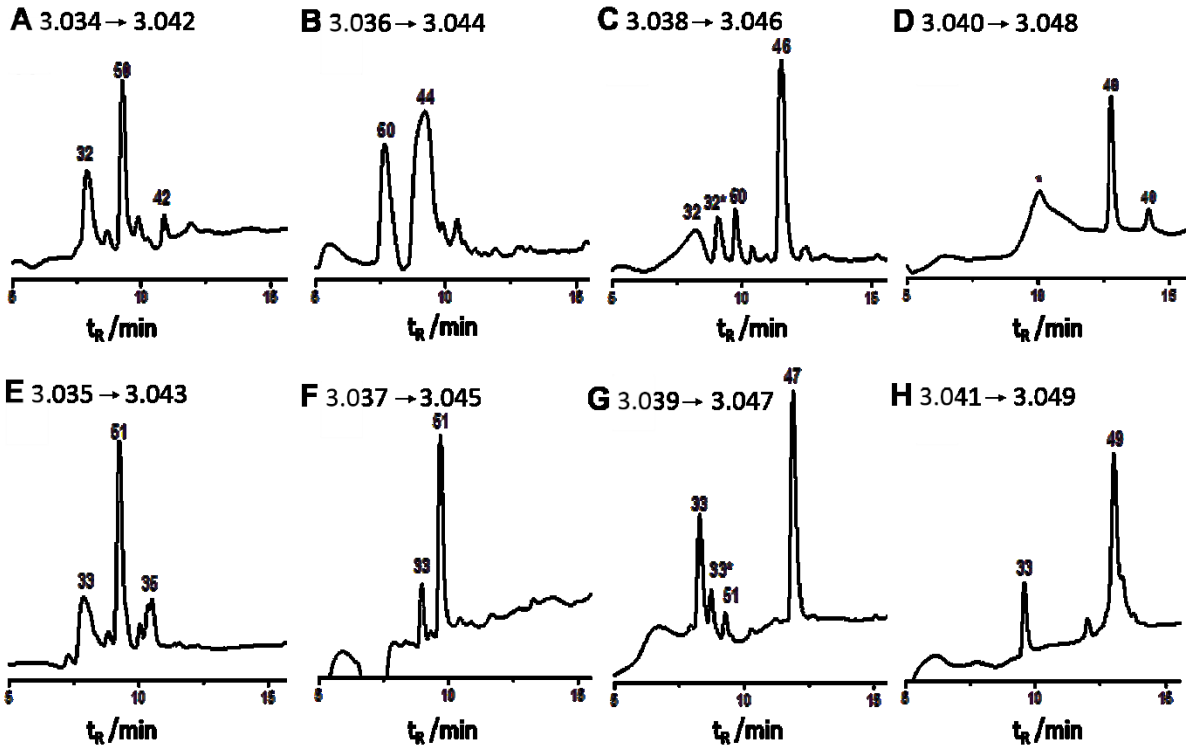
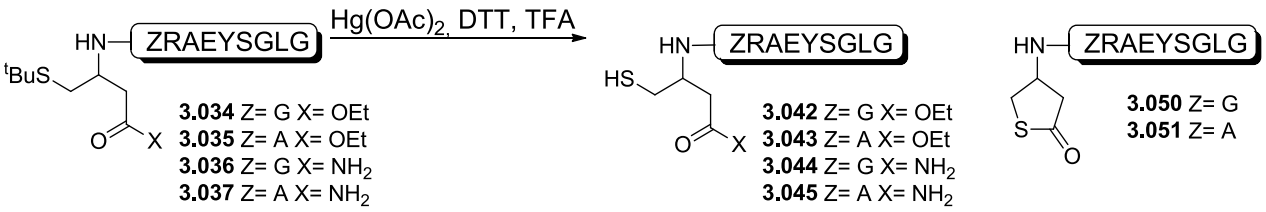




**Figure 5.1**: HPLC analysis of tert-butyl deprotection using method A to give unprotected auxiliary peptides **3.042-3.049**; peaks **3.032** and **3.033** correspond to peptides GRAYESGLG and ARAEYSGLG respectively; peak **3.032*** and **3.033*** correspond to N-acetylation of peptide **3.032** and **3.033** respectively; Gradient: 03-60% B in 20min, 210 nm.

**Removal of tert-Butyl using method B: Mercury Acetate**



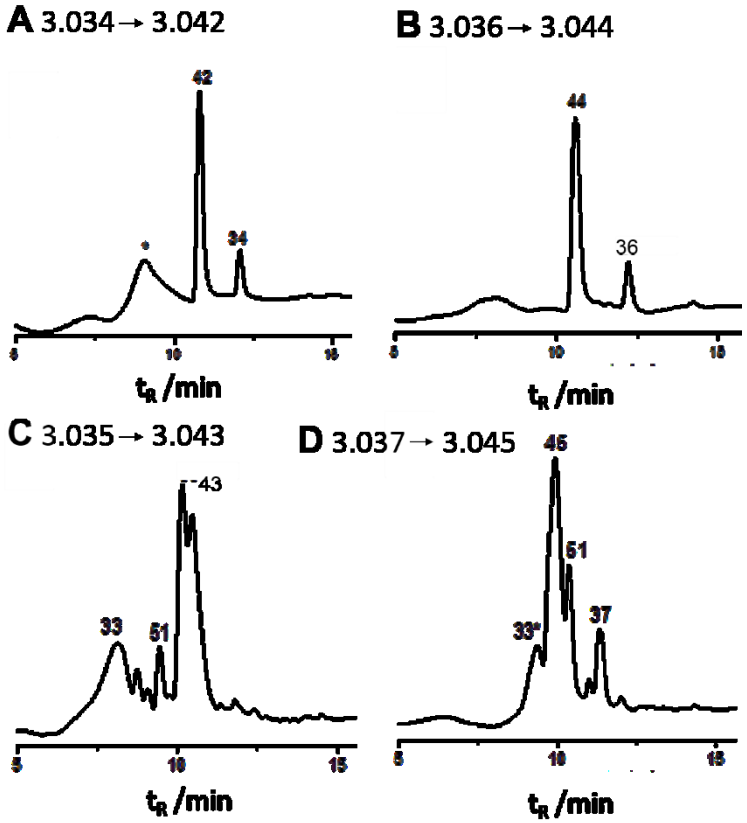




**Figure 5.2**: HPLC analysis of tert-butyl deprotection using method B to give unprotected auxiliary peptides **3.042-3.045**; peaks **3.032** and **3.033** correspond to peptides GRAYESGLG and ARAEYSGLG respectively; peak **3.32*** and **3.33*** correspond to N-acetylation of peptide **3.032** and **3.033** respectively; Gradient: 03-60% B in 20min, 210 nm.

Removal of tert-butyl group from peptide **3.034** to yield peptide **3.042** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 3.2 mg, 30%.

ESI-MS (positive mode): m/z = 1055.2 ($C_{46}H_{72}N_{13}O_{15}S$ $[M+H]^{+1}$, calcd.: 1055.2), 528.1 ($C_{46}H_{73}N_{13}O_{15}S$ $[M+2H]^{+2}$, calcd.: 528.1).

Removal of tert-butyl group from peptide **3.035** to yield peptide **3.043** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 1.5 mg, 14%.

ESI-MS (positive mode): m/z = 1069.2 ($C_{47}H_{74}N_{13}O_{15}S$ $[M+H]^{+1}$, calcd.: 1069.2), 535.1 ($C_{47}H_{75}N_{13}O_{15}S$ $[M+2H]^{+2}$, calcd.: 535.1).

Removal of tert-butyl group from peptide **3.036** to yield peptide **3.044** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 2.4 mg, 24%.
ESI-MS (positive mode): m/z = 1026.3 ($C_{42}H_{69}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1026.2), 513.6 ($C_{42}H_{70}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 513.6).

Removal of tert-butyl group from peptide **3.037** to yield peptide **3.045** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 1.9 mg, 18%.
ESI-MS (positive mode): m/z = 1040.2 ($C_{43}H_{71}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1040.2), 520.5 ($C_{43}H_{72}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 513.6).

Removal of tert-butyl group from peptide **3.038** to yield peptide **3.046** was carried out according to method A in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 4 mg, 36%.
ESI-MS (positive mode): m/z = 1102.4 ($C_{48}H_{73}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1102.3), 551.6 ($C_{48}H_{74}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 551.6).

Removal of tert-butyl group from peptide **3.039** to yield peptide **3.047** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 1.7 mg, 15%.
ESI-MS (positive mode): m/z = 1116.4 ($C_{49}H_{75}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1116.3), 558.7 ($C_{49}H_{76}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 558.7).

Removal of tert-butyl group from peptide **3.040** to yield peptide **3.048** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 1.8 mg, 16%.
ESI-MS (positive mode): m/z = 1116.2 ($C_{49}H_{75}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1116.3), 558.7 ($C_{49}H_{76}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 558.7).

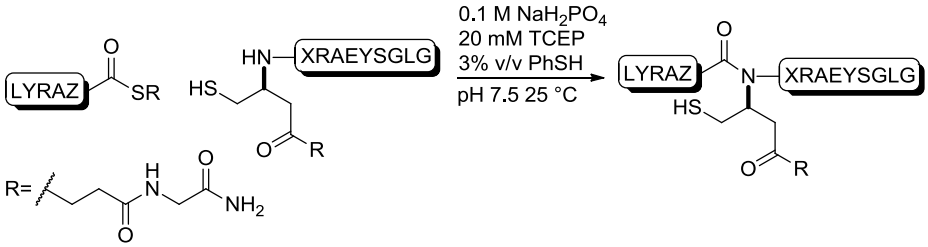
Removal of tert-butyl group from peptide **3.041** to yield peptide **3.049** was carried out according to method B in the general protocol. Purified using preparative HPLC. Gradient 03-60% B in 30 min. Isolated yield: 1.7 mg, 15%.

ESI-MS (positive mode): m/z = 1130.4 ($C_{50}H_{77}N_{14}O_{14}S$ $[M+H]^{+1}$, calcd.: 1130.3), 565.7 ($C_{50}H_{78}N_{14}O_{14}S$ $[M+2H]^{+2}$, calcd.: 565.7).

## 5.3.5 Auxiliary-Assisted Native Chemical Ligations:

Ligations were performed according to NCL method A as described in the general procedure.

**Table 5.1: Ligation yields**




**3.52** Z= G
**3.53** Z= A

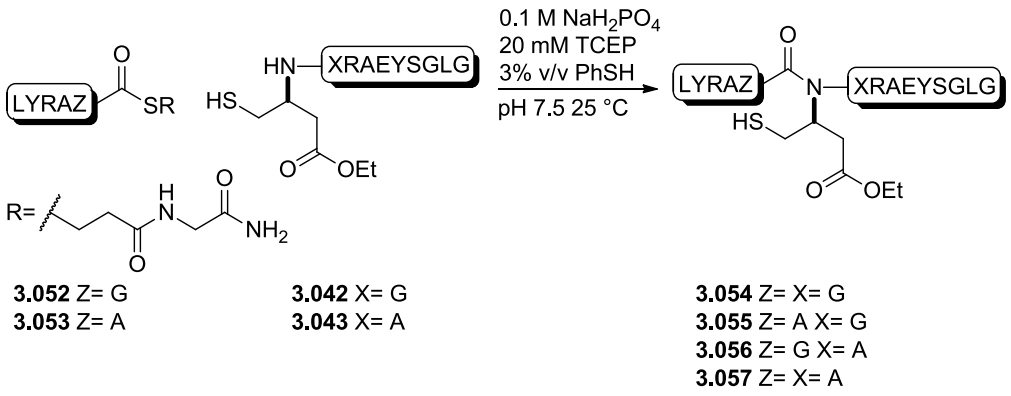
**3.042** X= G R= OEt
**3.043** X= A R= OEt
**3.044** X= G R= $NH_2$
**3.045** X= A R= $NH_2$
**3.046** X= G R= NHPh
**3.047** X= A R= NHPh
**3.048** X= G R= NMePh
**3.049** X= A R= NMePh

**3.054** Z= G X= G R= OEt
**3.055** Z= G X= A R= OEt
**3.056** Z= A X= G R= OEt
**3.057** Z= A X= A R= OEt
**3.058** Z= G X= G R= $NH_2$
**3.059** Z= G X= A R= $NH_2$
**3.060** Z= G X= G R= NHPh

**3.061** Z= G X= A R= NHPh
**3.062** Z= A X= G R= NHPh
**3.063** Z= A X= A R= NHPh
**3.064** Z= G X= G R= NMePh
**3.065** Z= G X= A R= NMePh
**3.066** Z= A X= G R= NMePh
**3.067** Z= A X= A R= NMePh

| TE (Z) | AP-(X-R) | LP | Yield[a] (t) | Isolated yield |
|---|---|---|---|---|
| **3.052 (G)** | **3.034** (G-OEt) | **3.054** | 99%(4 hr) | 56% |
| **3.052 (G)** | **3.035** (A-OEt) | **3.056** | 65% (8 hr) | 41% |
| **3.052 (G)** | **3.036** (G-$NH_2$) | **3.058** | 50% (2 hr) | ND |
| **3.052 (G)** | **3.038** (G-NHPh) | **3.060** | 905 (4 hr) | 57% |
| **3.052 (G)** | **3.039** (A-NHPh) | **3.062** | 80% (4 hr) | 37% |
| **3.052 (G)** | **3.040** (G-NMePh) | **3.064** | 37% (24 hr) | ND |
| **3.052 (G)** | **3.041** (A-NMePh) | **3.065** | 27% (24 hr) | ND |
| **3.053 (A)** | **3.034** (G-OEt) | **3.055** | 50% (8 hr) | 10 % |
| **3.053 (A)** | **3.035** (A-OEt) | **3.057** | ND | ND |
| **3.053 (A)** | **3.036** (G-$NH_2$) | **3.059** | 28% (4 hr) | ND |
| **3.053 (A)** | **3.038** (G-NHPh) | **3.061** | 80% (8 hr) | 17% |
| **3.053 (A)** | **3.039** (A-NHPh) | **3.063** | ND | ND |
| **3.053 (A)** | **3.040** (G-NMePh) | **3.065** | 14% (24 hr) | ND |
| **3.053 (A)** | **3.041** (A-NMePh) | **3.067** | ND | ND |

[a]yields were calculated by means of HPLC based on ratios of peak areas.



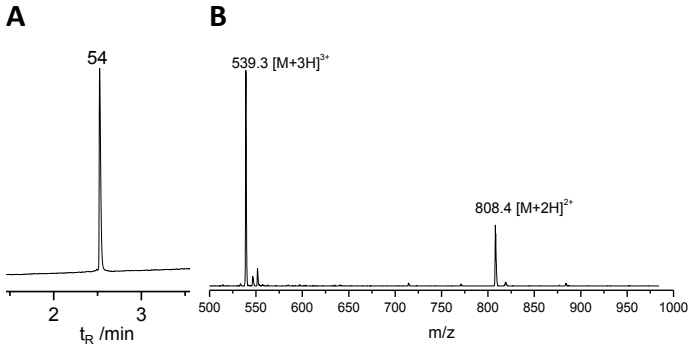




**Figure 5.3**: A) UPLC chromatogram of pure ligation product 3.054; B) ESI-MS of ligation product 3.054; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 808.4 ($C_{70}H_{112}N_{21}O_{21}S$ $[M+2H]^{+2}$, calcd.: 808.4), 539.3 ($C_{70}H_{113}N_{21}O_{21}S$ $[M+3H]^{+3}$, calcd.: 539.3).

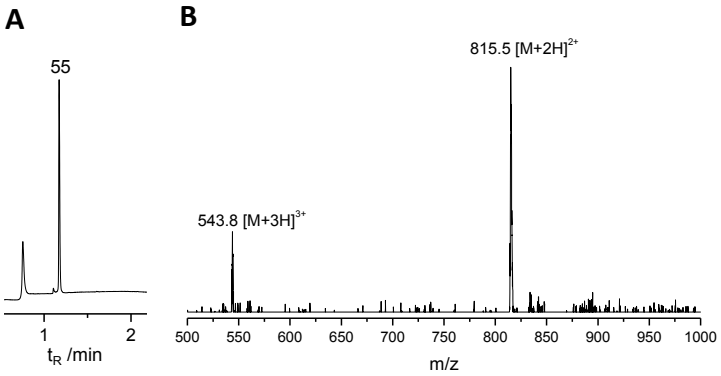



**Figure 5.4**: A) UPLC chromatogram of purified ligation product **3.055** showing spontaneous formation of thiolactone **3.050** from ligation product **3.055**;B) ESI-MS of ligation product **3.055**; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 815.4 ($C_{71}H_{114}N_{21}O_{21}S$ $[M+2H]^{+2}$, calcd.: 815.4), 543.9 ($C_{70}H_{115}N_{21}O_{21}S$ $[M+3H]^{+3}$, calcd.: 543.9).

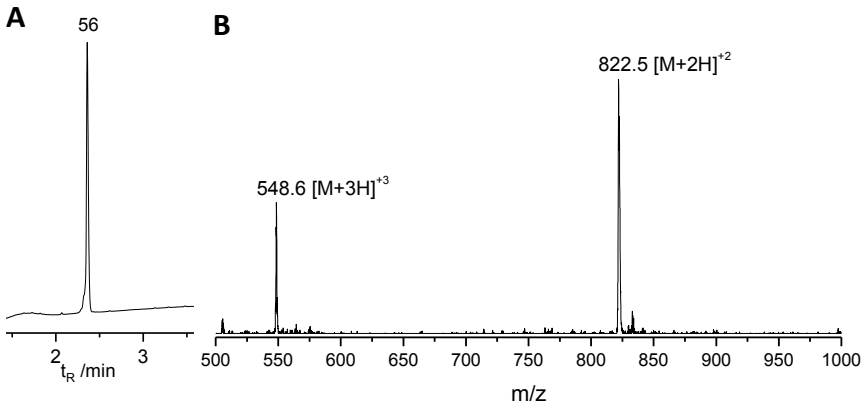



**Figure 5.5**: A) UPLC chromatogram of pure ligation product **3.056**;B) ESI-MS of ligation product **3.056**; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 837.4 ($C_{71}H_{113}N_{21}O_{21}SNa_2$ $[M+2Na]^{+2}$, calcd.: 837.4),815.4 ($C_{71}H_{114}N_{21}O_{21}S$ $[M+2H]^{+2}$, calcd.: 815.4), 543.9 ($C_{70}H_{115}N_{21}O_{21}S$ $[M+3H]^{+3}$, calcd.: 543.9).

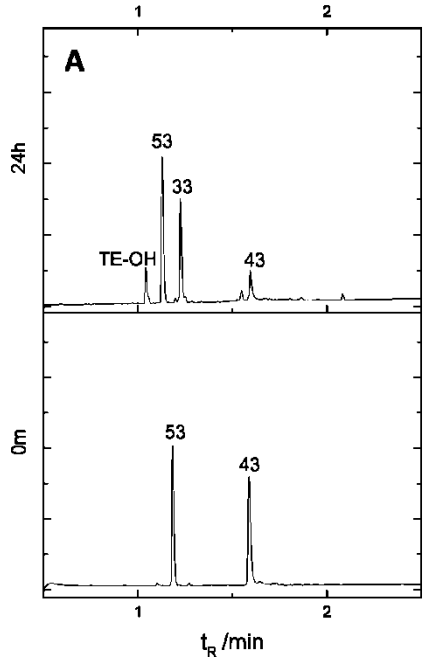



**Figure 5.6**: Ligation between thioester **3.053** and peptide auxiliary **3.043** at 0 hr and 24 hr; 03-60% B in 6 min, 210 nm.

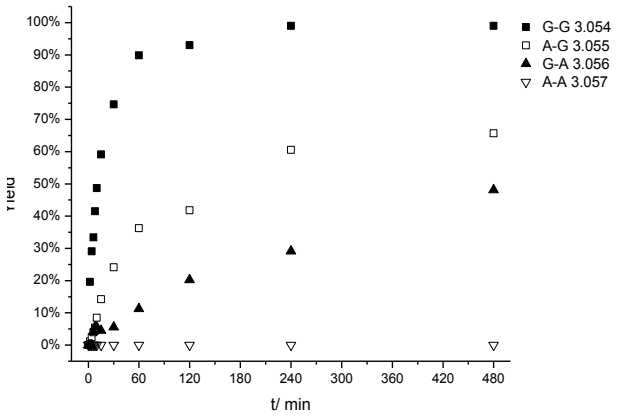



**Figure 5.7**: Time vs. yield for products **3.054-3.057**.

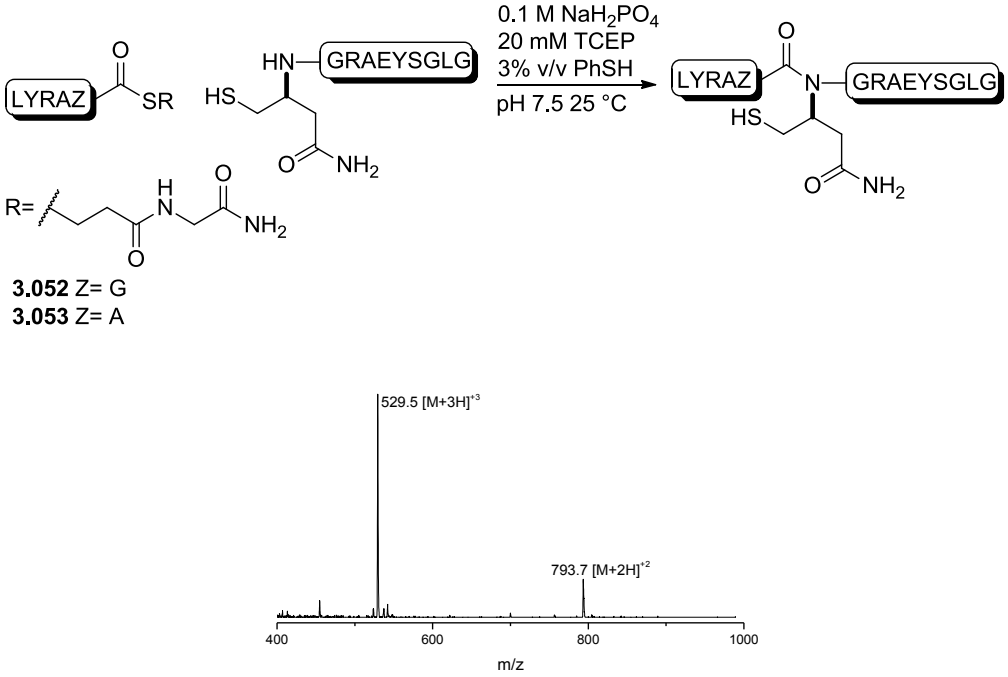






**Figure 5.8**: ESI-MS of ligation product **3.058**.

ESI-MS (positive mode): m/z = 793.7 ($C_{68}H_{109}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 793.9), 529.5 ($C_{68}H_{110}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 529.6).

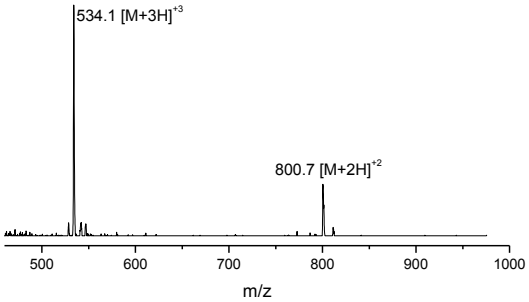



**Figure 5.9**: ESI-MS of ligation product **3.059**.

ESI-MS (positive mode): m/z = 800.7 ($C_{69}H_{111}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 800.9), 534.1 ($C_{69}H_{113}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 534.3).

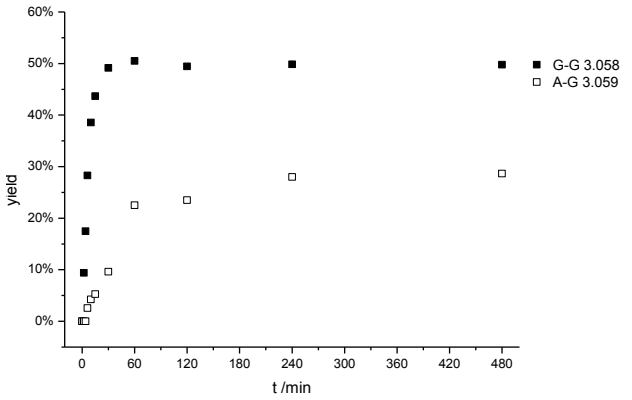
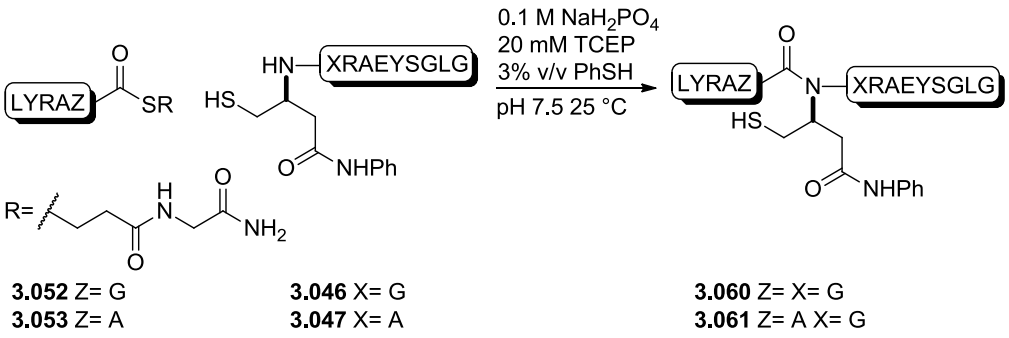



**Figure 5.10**: Time vs. yield for products **3.058, 3.059**.



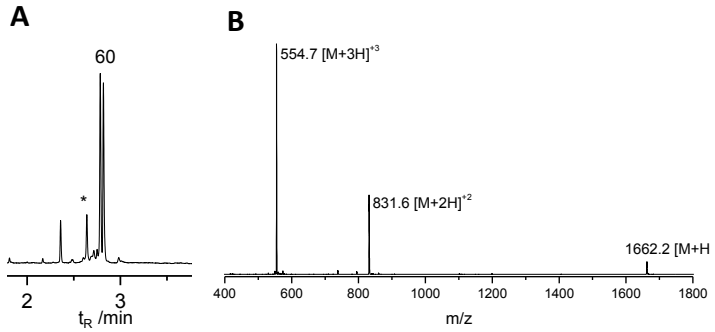




**Figure 5.11**: A) UPLC chromatogram of purified ligation product **3.060**;B) ESI-MS of ligation product **3.060**; Peak * corresponds to unidentified masses; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 1662.2 ($C_{74}H_{113}N_{22}O_{20}S$ $[M+H]^{+1}$, calcd.: 1662.9), 831.6 ($C_{74}H_{114}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 832.0) , 554.7 ($C_{74}H_{115}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 555.0).

.

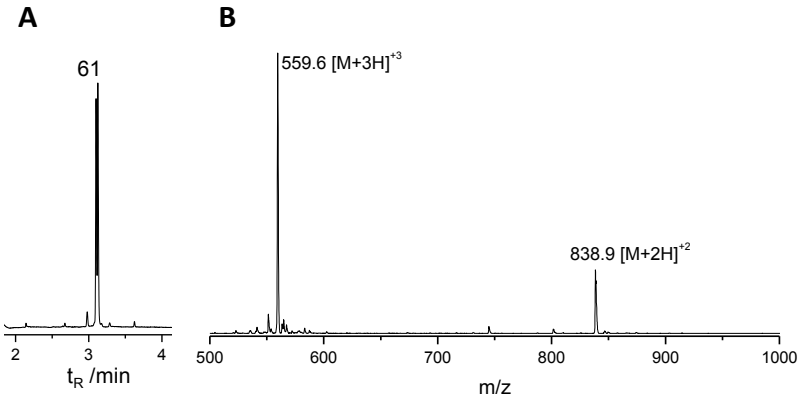



**Figure 5.12**: A) UPLC chromatogram of purified ligation product **3.061**;B) ESI-MS of ligation product **3.061**; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 838.9 ($C_{75}H_{115}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 839.0), 559.6 ($C_{75}H_{116}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 559.7).

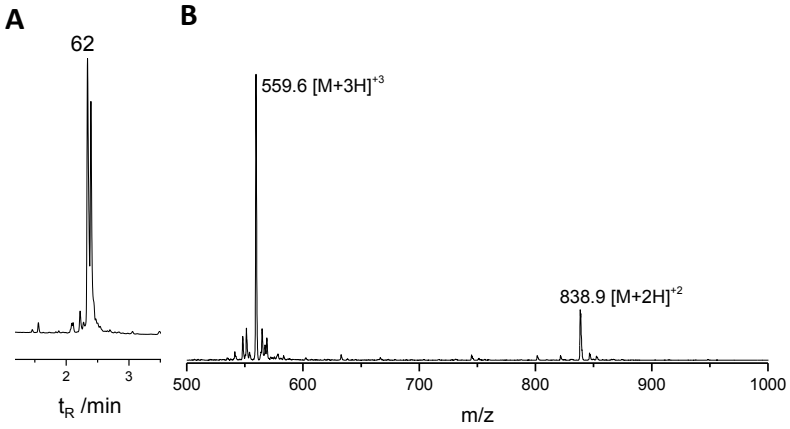



**Figure 5.13**: A) UPLC chromatogram of purified ligation product **3.062**;B) ESI-MS of ligation product **3.062**; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 838.8 ($C_{75}H_{115}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 839.0), 559.6 ($C_{75}H_{116}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 559.7).

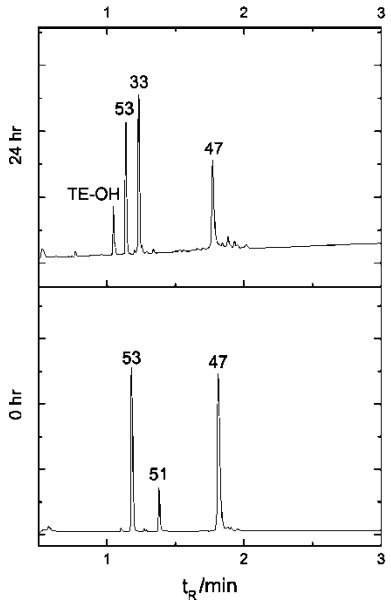



**Figure 5.14**: Ligation between thioester **3.053** and peptide auxiliary **3.047** at 0 hr and 24 hr; 03-60% B in 6 min, 210 nm.

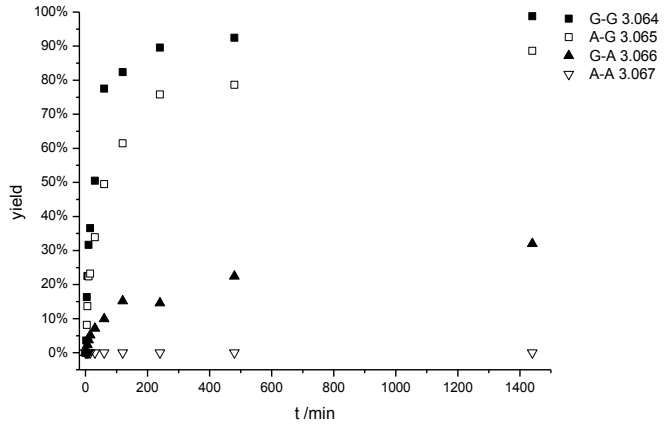
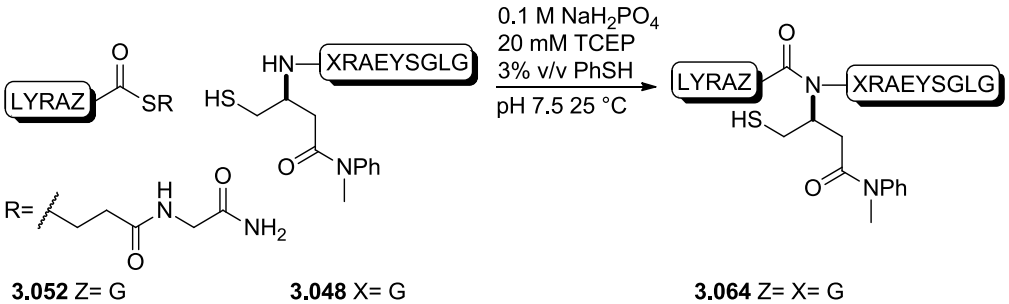



**Figure 5.15**: Time vs. yield for products **3.060- 3.063**.




**3.052** Z= G
**3.053** Z= A

**3.048** X= G
**3.049** X= A

**3.064** Z= X= G
**3.065** Z= A X= G
**3.066** Z= G X= A
**3.067** Z= X= A

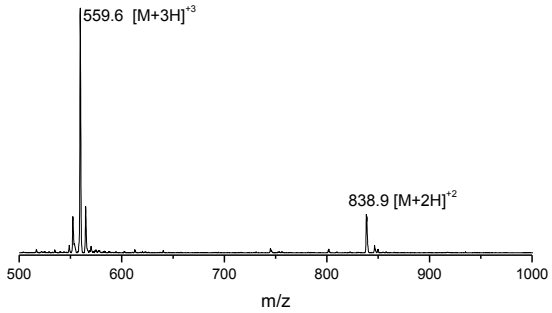



**Figure 5.16**: ESI-MS of ligation product **3.064**.

ESI-MS (positive mode): m/z = 838.9 ($C_{75}H_{115}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 839.0), 559.6 ($C_{75}H_{116}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 559.7).

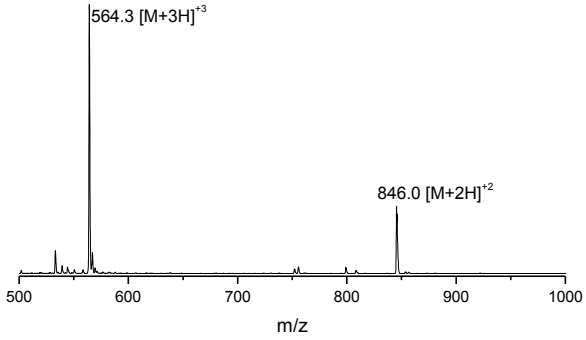



**Figure 5.17**: Ligation between thioester **3.053** and peptide auxiliary **3.048** to give ligation product **3.065** at 0 hr and 24 hr. Peak * correspond to unidentified masses; 03-60% B in 6 min, 210 nm.

ESI-MS (positive mode): m/z = 846.0 ($C_{76}H_{117}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 847.0), 564.3 ($C_{76}H_{118}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 565.0).

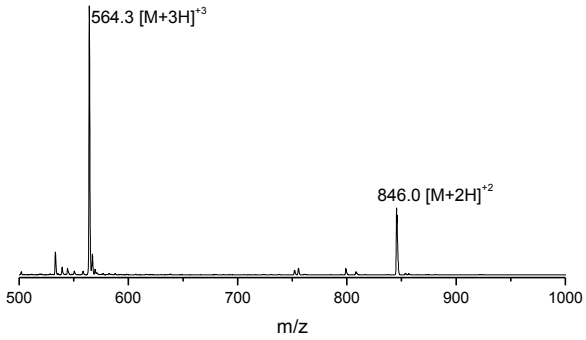



**Figure 5.18**: ESI-MS of ligation product **3.066**.

ESI-MS (positive mode): m/z = 846.9 ($C_{76}H_{117}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 846.0), 564.3 ($C_{76}H_{118}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 564.3).

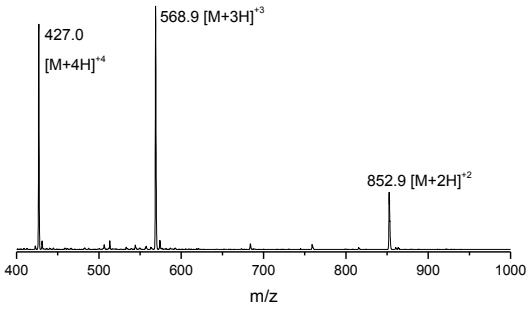



**Figure 5.19**: ESI-MS of ligation intermediate **3.067***

ESI-MS (positive mode): m/z = 852.9 ($C_{76}H_{117}N_{22}O_{20}S$ $[M+2H]^{+2}$, calcd.: 853.0), 568.9 ($C_{76}H_{118}N_{22}O_{20}S$ $[M+3H]^{+3}$, calcd.: 569.0) , 427.0 ($C_{76}H_{119}N_{22}O_{20}S$ $[M+4H]^{+4}$, calcd.: 427.0).

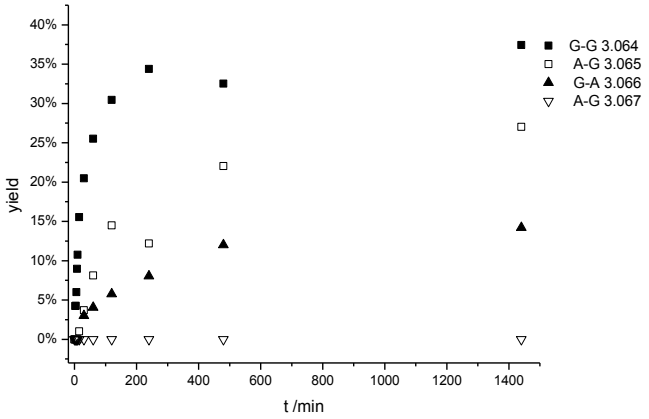



**Figure 5.20**: Time vs. yield for products **3.064- 3.067**.

## 5.3.6 Removal of Auxiliary

Auxiliary removal was performed as described in method A in the general protocol.

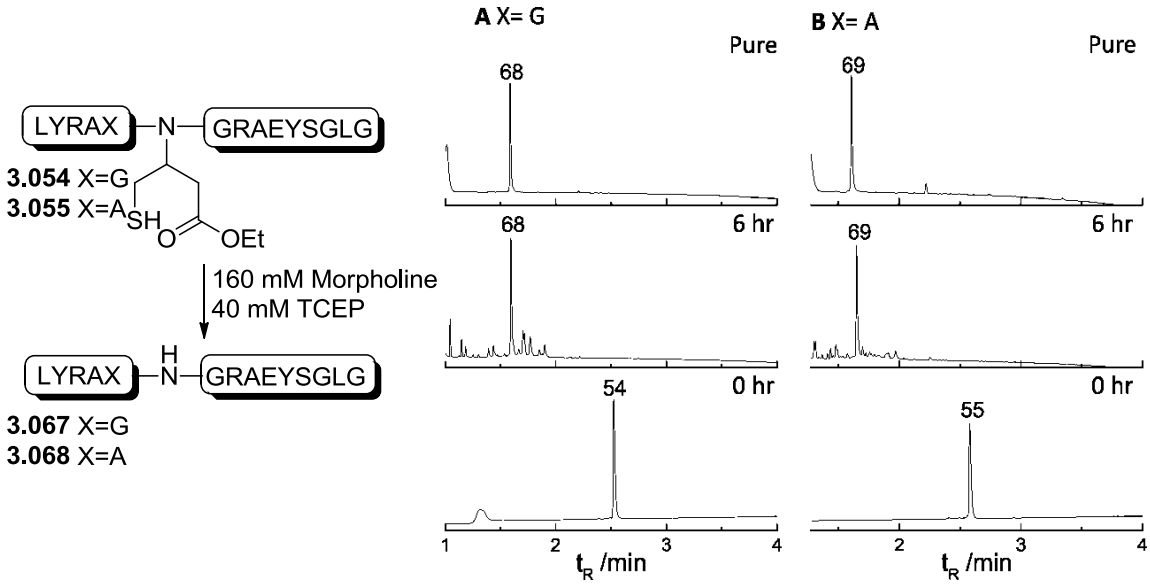



**Figure 5.21**: UPLC analysis showing the removal of auxiliary from ligation product **3.054** to give native peptide **3.068** (**A**) and from ligation product **3.055** to give native peptide **3.069** (**B**) at 0 and 6 hr; Gradient: 03-60% B in 4 min, 210 nm.

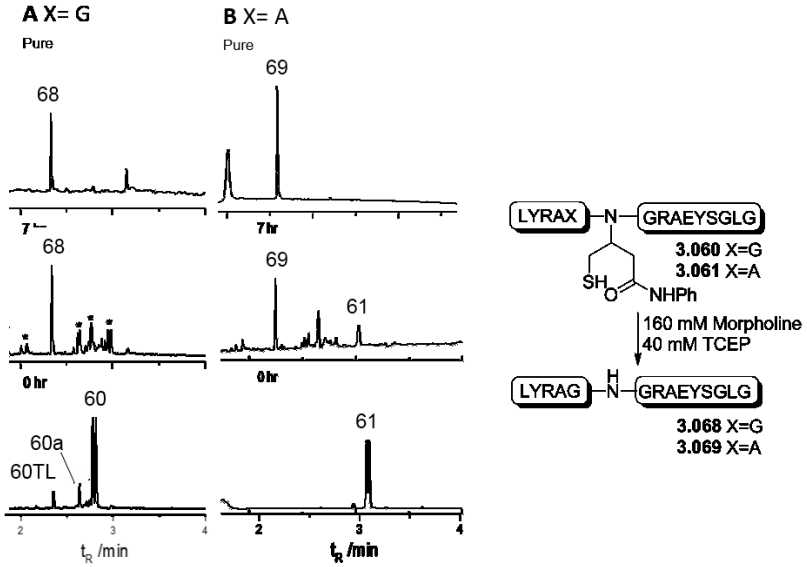



**Figure 5.22**: UPLC analysis showing the removal of auxiliary from ligation product **3.060** to give native peptide **3.068** (**A**) and from ligation product **3.061** to give native peptide **3.069** (**B**) at 0 and 7 hr; Gradient: 03-60% B in 4 min, 210 nm.

## 5.3.7 Total synthesis of DCD-1L

**Ligation and Auxiliary Removal**: Ligation was performed using method A as described in the general protocol in 1.2 μmol scale.

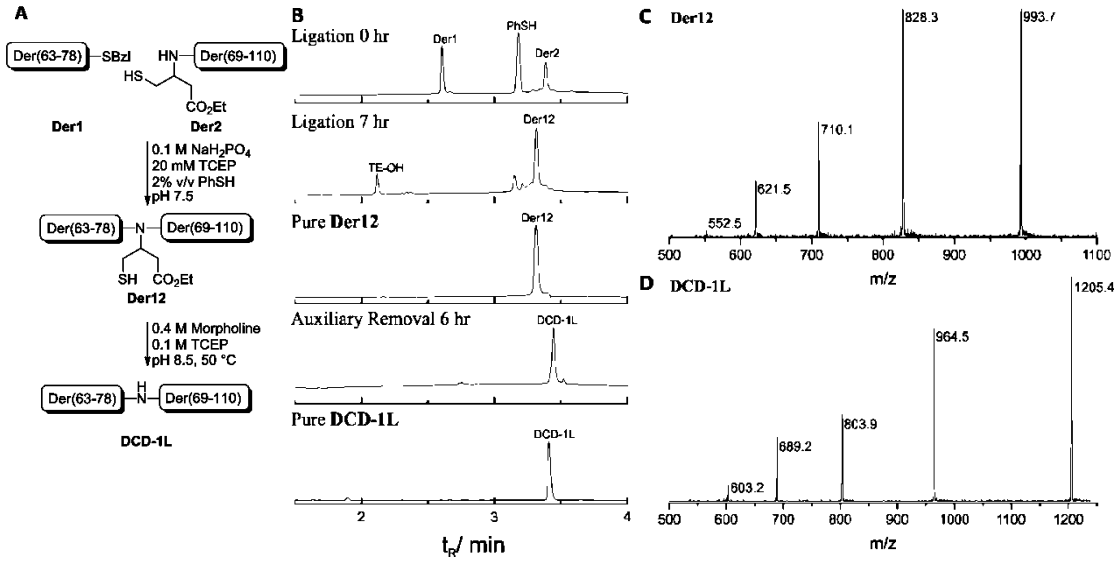



**Figure 5.23**: Synthesis of DCD-1L. **A)** Scheme depicting synthetic route; **B)** UPLC analysis of ligation between **Der1** and **Der2** and auxiliary removal to yield native **DCD-1L**; **C)** MS analysis **Der12** with observed mass 4963.6 Da (calc. 4963.7 Da); **D)** MS analysis of **DCD-1L** with observed mass 4817.4 Da (calc. 4817.5 Da); 03-70% B in 6 min, 210 nm.

**Der12** H-Der(63-78)-Aux-Der(69-110)-$NH_2$ 3.0 mg, 4963.7 gr/mol, 0.60 μmol, 50%

ESI-MS (positive mode): m/z = 993.7 ($C_{216}H_{379}N_{58}O_{72}S$ $[M+5H]^{+5}$, calcd.: 993.8) , 828.3 ($C_{216}H_{380}N_{58}O_{72}S$ $[M+6H]^{+6}$, calcd.: 828.3), 710.1 ($C_{216}H_{381}N_{58}O_{72}S$ $[M+7H]^{+7}$, calcd.: 710.1), 621.5 ($C_{216}H_{382}N_{58}O_{72}S$ $[M+8H]^{+8}$, calcd.: 621.5), 552.5 ($C_{216}H_{382}N_{58}O_{72}S$ $[M+9H]^{+9}$, calcd.: 552.5).

**DCD-1L** H-Der(63-110)-$NH_2$, 1.2 mg, 4817.5 gr/mol, 0.24 μmol, 40%

ESI-MS (positive mode): m/z = 1205.4 ($C_{210}H_{364}N_{58}O_{70}$ $[M+4H]^{+4}$, calcd.: 1205.4), 964.5 ($C_{210}H_{365}N_{58}O_{70}$ $[M+5H]^{+5}$, calcd.: 964.5) , 803.9 ($C_{210}H_{366}N_{58}O_{70}$ $[M+6H]^{+6}$, calcd.: 803.9), 689.2 ($C_{210}H_{367}N_{58}O_{70}$ $[M+7H]^{+7}$, calcd.: 689.2), 603.2 ($C_{210}H_{368}N_{58}O_{70}$ $[M+8H]^{+8}$, calcd.: 603.2).

**Stability of peptides to long exposure of auxiliary removal buffer**

To test the protein stability under the removal conditions pure Dermicidin-1L **DCD-1L** was redissolved in the removal buffer and shaken for 24 hr with aliquots taken at 3, 6 and 24 hr. results indicated that the peptide is stable for prolonged treatment with small side-products appearing after 24 hr. The side-products showed same mass as the original protein probably

hinting towards racemization of different residues after the long treatment at an elevated temperature and pH (figure 5.24).

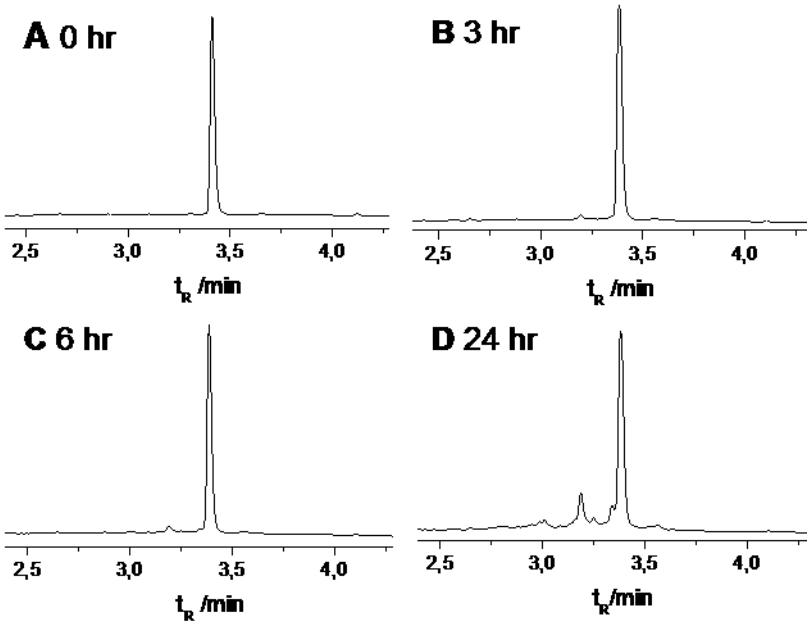



**Figure 5.24**: UPLC analysis of pure **Der-1L** in auxiliary removal buffer over time. **A)** 0 hr, **B)** 3 hr, **C)** 6hr, **D)** 24 hr, 03-70%B in 6 min, 210 nm.

# 5.4 Experimental Data for Auxiliary-Assisted SPCL

## 5.4.1 Synthesis of Auxiliary Precursor 2-((bis(4-methoxyphenyl) (phenyl) methyl)thio)-2-phenylacetaldehyde (3.072)

The auxiliary **3.072** was synthesized as described in reference [95].

## 5.4.2 SPPS, Preparation of peptide building blocks and Chromatographic Purification of Proteins and Peptides

**Synthesis of Peptides Aux-MUC1(1-20)-His$_6$ (3.079), Aux-MUC1(1-20)-Dbz-Gly (3.086) and Aux-MUC1(41-80)-His$_6$ (3.114)**

*Sequence of MUC1(1-20):* GVTSAPDTRPAPGSTAPPAH

*SPPS*: The peptides were synthesized as described in the general protocol on a Rink Amide resin (0.19 mmol/g) in a 10 μmol scale.

Aux-MUC1(1-20)-Dbz-Gly **3.086** was synthesized in a 10 μmol scale as described in the general procedure for synthesis of Dbz peptides.

*Reductive amination*: was performed as described in method C in the general protocol.

*Cleavage of protecting groups and isolation of peptide*: Performed as described previously under the general procedure for cleavage and purification of peptides.

**3.079 Aux-MUC1(1-20)-His$_6$** 12.4 mg, 3101.0 gr/mol, 4 μmol, 40%, %,
ESI-MS (positive mode): m/z = 1034.3 ($C_{124}H_{179}N_{44}O_{33}S$ $[M+3H]^{+3}$, calcd.: 1034.3), 776.2 ($C_{124}H_{180}N_{44}O_{33}S$ $[M+4H]^{+4}$, calcd.: 776.0) , 621.4 ($C_{124}H_{181}N_{44}O_{33}S$ $[M+5H]^{+5}$, calcd.: 621.0), 517.9 ($C_{124}H_{182}N_{44}O_{33}S$ $[M+6H]^{+6}$, calcd.: 517.7), 444.0 ($C_{124}H_{182}N_{44}O_{33}S$ $[M+7H]^{+7}$, calcd.: 443.9).

**3.086 Aux-MUC1(1-20)-Dbz-Gly**, 7.4 mg, 2468.7 gr/mol ,3 μmol, 30%,
ESI-MS (positive mode) m/z = 823.9 ($C_{88}H_{139}N_{29}O_{29}$ $[M+3H]^{3+}$, calcd: 823.8), 618.4 ($C_{88}H_{140}N_{29}O_{29}$ $[M+4H]^{4+}$, calcd: 618.1), 494.0 ($C_{88}H_{141}N_{29}O_{29}$ $[M+5H]^{5+}$, calcd. 494.7).

**3.114 Aux-MUC1(41-80)-$His_6$** 7.0 mg, 4715 gr/mol, 1.5 µmol, 15%.

ESI-MS (positive mode) m/z = 1179.5 ($C_{204}H_{307}N_{69}O_{60}S$ $[M+4H]^{4+}$, calcd. 1179.5), 943.8 ($C_{204}H_{308}N_{69}O_{60}S$ $[M+5H]^{5+}$, calcd. 943.8) 786.7 ($C_{204}H_{309}N_{69}O_{60}S$ $[M+6H]^{6+}$, calcd. 786.7), 674.5 ($C_{204}H_{310}N_{69}O_{60}S$ $[M+7H]^{7+}$, calcd. 674.4), 590.3 ($C_{204}H_{311}N_{69}O_{60}S$ $[M+8H]^{8+}$, calcd. 590.3).

**Synthesis of thioester peptide $His_6$-MUC1(1-20)-SBzl (3.093) and H-MUC1(1-40)-SBzl (3.094) and $His_5$-MUC1(1-20)-SBzl (3.102) using safety-catch linker approach**

Peptides were synthesized as described in the general protocol in 10 µmol scale.

**3.093** $His_6$-MUC1(1-20)-SBzl, 13.8 mg, 3071.6 gr/mol, 4.5 µmol, 45%.

ESI-MS (positive mode) m/z = 1024.6 ($C_{123}H_{178}N_{43}O_{34}S$ $[M+3H]^{3+}$, calcd. 1024.7), 768.9 ($C_{123}H_{179}N_{43}O_{34}S$ $[M+4H]^{4+}$, calcd. 768.9) 615.4 ($C_{123}H_{180}N_{43}O_{34}S$ $[M+5H]^{5+}$, calcd. 615.2), 512.8 ($C_{123}H_{181}N_{43}O_{34}S$ $[M+6H]^{6+}$, calcd. 512.8), 439.7 ($C_{123}H_{182}N_{43}O_{34}S$ $[M+7H]^{7+}$, calcd. 439.7).

**3.094** H-MUC1(1-40)-SBzl 13.5 mg, 3862 gr/mol, 3.5 µmol, 35%.

ESI-MS (positive mode) m/z = 966.6 ($C_{167}H_{262}N_{50}O_{55}S$ $[M+4H]^{4+}$, calcd. 966.6), 773.2 ($C_{167}H_{263}N_{50}O_{55}S$ $[M+5H]^{5+}$, calcd. 773.5), 644.5 ($C_{167}H_{264}N_{50}O_{55}S$ $[M+6H]^{6+}$, calcd. 644.7), 552.6 ($C_{167}H_{265}N_{50}O_{55}S$ $[M+7H]^{7+}$, calcd. 552.7).

**3.102** $His_5$-MUC1(1-20)-SBzl,13.4 mg, 2678.9 gr/mol, 5.0 µmol, 50%.

ESI-MS (positive mode) m/z = 894.0 ($C_{117}H_{171}N_{40}O_{33}S$ $[M+3H]^{3+}$, calcd. 894.0), 670.7 ($C_{117}H_{172}N_{40}O_{33}S$ $[M+4H]^{4+}$, calcd. 670.7), 536.8 ($C_{117}H_{173}N_{40}O_{33}S$ $[M+5H]^{5+}$, calcd. 536.8).

**Synthesis of Hydrazide Peptide $His_6$-MUC1(1-20)-$NHNH_2$ (3.111)**

*Synthesis of hydrazine 2-chlorotrytyl resin*: Resin (325 mg, 1.3 mmol/gr, 0.5 mmol) swelled in DMF. A mixture of 1:1:1 v/v $NEt_3$: $H_2O$: DMF (1 ml) added for 60 min. and MeOH (1 ml) is added for 30 min. Resin is washed with MeOH+ 5% v/v DIPEA, MeOH, DMF and DCM (X 3 each) and swelled in DMF.

*Loading of first amino acid*: Fmoc-His(Trt)-OH (31 mg, 0.05 mmol, 5 eq), HCTU (20.7 mg, 0.05 mmol), HOBt (6.8 mg, 0.05 mmol) and NMM (11 µl, 0.1 mmol, 10 eq) in 125 µl DMF is added to resin and coupled 2 X 30 min.

*SPPS*: Rest of peptide was synthesized on an automated peptide synthesizer as described in the general protocol.

*Cleavage of protecting groups and isolation of peptide*: Performed as described previously under the general procedure for cleavage and purification of peptides.

**3.111** $His_6$-MUC1(1-20)-$NHNH_2$ 15 mg, 2723.9 gr/mol, 5.5 μmol, 55%.
ESI-MS (positive mode) m/z = 908.8 ($C_{110}H_{167}N_{42}O_{32}$ $[M+3H]^{3+}$, calcd. 909.0), 682.7 ($C_{110}H_{168}N_{42}O_{32}$ $[M+4H]^{4+}$, calcd. 682.0), 545.8 ($C_{110}H_{169}N_{42}O_{32}$ $[M+5H]^{5+}$, calcd. 545.8).

## 5.4.3 Auxiliary-Assisted SPCL in N→C and C→N direction

**Ligation between His6-MUC1(1-20)-SBzl (3.093) and Aux-MUC1(1-20)-Dbz-Gly (3.086)***:* done as described in the general protocol. In 0.5 μmol scale to yield product **3.095**.
ESI-MS (positive mode) m/z = 1106.7 ($C_{213}H_{315}N_{73}O_{62}SNa_5$ $[M+5H]^{5+}$, calcd. 1107.0), 903.7 ($C_{213}H_{321}N_{73}O_{62}S$ $[M+6H]^{6+}$, calcd. 903.5), 774.7 ($C_{213}H_{322}N_{73}O_{62}S$ $[M+7H]^{7+}$, calcd. 774.6), 667.9 ($C_{213}H_{323}N_{73}O_{62}S$ $[M+8H]^{8+}$, calcd. 667.9), 602.9 ($C_{213}H_{343}N_{73}O_{62}S$ $[M+9H]^{9+}$, calcd. 602.7).

**Ligation between H-MUC1(1-40)-SBzl (3.094) and Aux-MUC1(1-20)-$His_6$ (3.079)**: done as described in the general protocol. In 0.5 μmol scale to yield product **3.096**. ESI-MS (m/z):
Observed Mass: 1140.7 $[M+6H]^{6+}$, 977.9 $[M+7H]^{7+}$, 856.0 $[M+8H]^{8+}$, 761.0 $[M+9H]^{9+}$, 685.3 $[M+10H]^{10+}$, 622.1 $[M+11H]^{11+}$ Calculated Mass:. 1140.7 $[M+6H]^{6+}$, 977.9 $[M+7H]^{7+}$, 855.8 $[M+8H]^{8+}$, 760.8 $[M+9H]^{9+}$, 684.8 $[M+10H]^{10+}$622.1 $[M+11H]^{11+}$.
ESI-MS (positive mode) m/z = 1140.7 ($C_{284}H_{434}N_{94}O_{87}S$ $[M+6H]^{6+}$, calcd. 1140.7), 977.9 ($C_{284}H_{435}N_{94}O_{87}S$ $[M+7H]^{7+}$, calcd. 977.9), 856.0 ($C_{284}H_{436}N_{94}O_{87}S$ $[M+8H]^{8+}$, calcd. 856.0), 761.0 ($C_{284}H_{437}N_{94}O_{87}S$ $[M+9H]^{9+}$, calcd. 760.8), 685.3 ($C_{284}H_{439}N_{94}O_{87}S$ $[M+10H]^{90+}$, calcd. 684.8), 622.1 ($C_{284}H_{438}N_{94}O_{87}S$ $[M+11H]^{11+}$, calcd. 622.1).

**Removal of Auxiliary from Ligation Product (3.095)**: done according to general protocol method C in 0.1 μmol scale to yield His6-MUC1(1-40)-Dbz-Gly **3.097**.
ESI-MS (positive mode) m/z = 1079.8 ($C_{276}H_{420}N_{94}O_{87}Na_5$ $[M+5H]^{5+}$, calcd. 1079.8), 880.8 ($C_{276}H_{426}N_{94}O_{87}$ $[M+6H]^{6+}$, calcd. 880.9), 755.4 ($C_{276}H_{427}N_{94}O_{87}$ $[M+7H]^{7+}$, calcd. 755.2), 661.3 ($C_{276}H_{428}N_{94}O_{87}$ $[M+8H]^{8+}$, calcd. 660.9), 587.8 ($C_{276}H_{429}N_{94}O_{87}$ $[M+9H]^{9+}$, calcd. 587.6).

### 5.4.3.1 Stability Study: His-Tag peptide binding to Ni-NTA resin

The resin bound peptide was treated with ligation buffer containing 20 mM TCEP and 10 mM MPAA to a concentration of 10 mM and kept at rt. Aliquots from supernatant were taken and tested for peptide presence. After 4 days, no peptide could be detected in the supernatant.

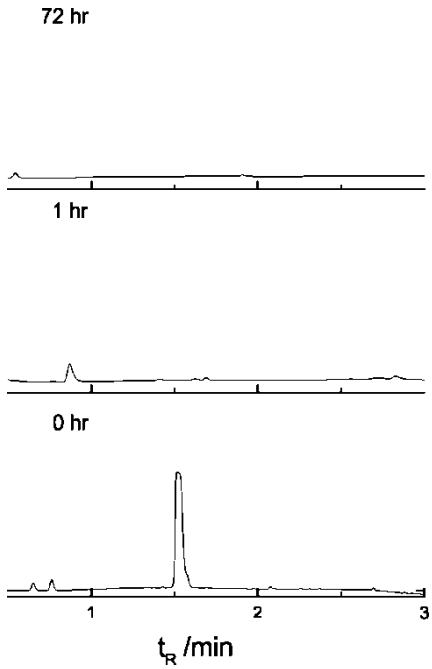



**Figure 5.25**: Stability test of Ni-His-Tag binding after 72 hr. 03-60% B in 6 min, 210 nm.

## 5.4.4 Activation of Latent Thioesters for Sequential Auxiliary-Assisted SPCL

**Transthioesterification of $His_5$-MUC1(1-20)-SBzl (3.102) to yield $His_5$-MUC1(1-20)-SPh (3.105b) or $His_5$-MUC1(1-20)- MESA (3.105c)**: done according to method B in the general protocol.

**3.105b** $His_5$-MUC1(1-20)-SPh:
ESI-MS (positive mode) m/z = 889.4 ($C_{116}H_{169}N_{40}O_{31}S$ $[M+3H]^{3+}$, calcd. 889.3), 667.5 ($C_{116}H_{170}N_{40}O_{31}S$ $[M+4H]^{4+}$, calcd. 667.2), 533.9 ($C_{116}H_{171}N_{40}O_{31}S$ $[M+5H]^{5+}$, calcd. 534.0).
**3.105c** $His_5$-MUC1(1-20)-MESA
ESI-MS (positive mode) m/z = 899.6 ($C_{112}H_{169}N_{40}O_{34}S_2$ $[M+3H]^{3+}$, calcd. 900.0), 675.3 ($C_{112}H_{169}N_{40}O_{34}S_2$ $[M+4H]^{4+}$, calcd. 675.2), 540.4 ($C_{112}H_{169}N_{40}O_{34}S_2$ $[M+5H]^{5+}$, calcd. 540.0).

**Converting $His_6$-MUC1(1-20)-$NHNH_2$ (3.111) to thioester $His_6$-MUC1(1-20)-MPAA (3.105a)**: done according to Method A in the general procedure.

**3.105a** $His_6$-MUC1(1-20)-MPAA:
ESI-MS (positive mode) m/z = 954.1 ($C_{118}H_{171}N_{40}O_{33}S$ $[M+3H]^{3+}$, calcd. 954.4), 715.9 ($C_{118}H_{172}N_{40}O_{33}S$ $[M+4H]^{4+}$, calcd. 716.0), 573.0 ($C_{118}H_{173}N_{40}O_{33}S$ $[M+5H]^{5+}$, calcd. 573.0), 477.7 ($C_{118}H_{174}N_{40}O_{33}S$ $[M+6H]^{6+}$, calcd. 477.7).

## 5.4.5 Total Chemical Synthesis of MUC1(1-80)-$His_6$

*Loading on His-Tag Ni-NTA affinity resin*: The Ni-NTA resin (70 µl, 70 mg peptide/ml, 50% solution) in 30% ethanol in water is taken, washed with degassed deionized water and swelled in a freshly degassed buffer of 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP pH 7.2 for 20 min. Peptide Aux-MUC1-$His_6$ **3.114** (2 mg, 4715 gr/mol, 0.42 µmol) is dissolved in same buffer (80 µl, 5 mM) and added to resin. After 1 hr, UPLC analysis of the supernatant showed a complete anchoring of the peptide to the resin.

*Ligation of Aux-MUC1(41-80)-$His_6$ **3.114** and H-MUC1(1-40)-SBzl **3.112** to give H-MUC1(1-40)-Aux-MUC1(41-80)-$His_6$ **3.115***: H-MUC1(1-40)-SBzl **3.112** (2 mg, 3862 g/mol, 0.50 µmol, 1.2 eq) was dissolved in 40 µl freshly degassed buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP, 10 mM MPAA pH 7.2.and added to resin. Mixture was incubated at 25 °C for 72 hr. Every 24 hr, resin was washed and a new batch of thioester H-MUC1(1-40)-SBzl **3.112** (0.8 mg, 0.21 µmol, 0.5 eq) in ligation buffer was added to resin.

*Monitoring the ligation*: Resin is washed with buffer of 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP pH 7.2 and degassed water. Analytical amount of resin is taken, and treated with 20 µl elution buffer containing 250 mM imidazole and 20 mM TCEP in water pH 8.0 for 20 min. 2 µl are taken and diluted with 18 µl $H_2O$+ 0.1% TFA and injected on UPLC-MS.

*Stability Control*: To check that the peptide anchoring to the resin is stable to the long exposure of TCEP and MPAA an aliquot from the supernatant was taken and injected on UPLC-MS. The results showed no presence of the auxiliary peptide in the supernatant. Only the thioester peptide **3.112** and the thioester MPAA exchange species were detected.

*Elution of ligation product H-MUC1(1-40)-Aux-MUC1(41-80)-$His_6$ **3.115** from solid support*: The resin is thoroughly washed with buffer containing 6M Gn.HCl, 0.1 M $NaH_2PO_4$, 20 mM TCEP pH 7.2, water, H2O+ 0.1% TFA, water. Resin is drained and the elution buffer (40 µl) is added for 30 min. The elution buffer is taken and process is repeated once more. The resin is then washed with elution buffer (2 X 10 µl) and all washes are combined.

ESI-MS (positive mode) m/z = 1208.3 ($C_{364}H_{560}N_{119}O_{114}S$ $[M+7H]^{7+}$, calcd. 1208.5), 1057.2 ($C_{364}H_{561}N_{119}O_{114}S$ $[M+8H]^{8+}$, calcd. 1057.5, 940.5 ($C_{364}H_{562}N_{119}O_{114}S$ $[M+9H]^{9+}$, calcd.940.1), 846.5 ($C_{364}H_{563}N_{119}O_{114}S$ $[M+10H]^{10+}$, calcd. 846.2), 769.3 ($C_{364}H_{564}N_{119}O_{114}S$ $[M+11H]^{11+}$, calcd. 769.4), 705.4 ($C_{364}H_{565}N_{119}O_{114}S$ $[M+12H]^{12+}$, calcd. 705.4).

*Removal of Auxiliary from ligation product* H-*MUC1(1-40)-Aux-MUC1(41-80)-His$_6$* ***3.115*** *to yield H-MUC1(1-80)-His$_6$ peptide* ***3.116***: To an Eppendorf containing ligation product 4 in the elution buffer (100 µl) is added 100 µl of aqueous buffer containing 400 mM morpholine and 100 mM TCEP pH 8.5 and incubated at 25 °C for 24 hr after which a full conversion to peptide **3.116** is achieved as observed by UPLC-MS.

*Purification of* H-*MUC1(1-80)-His$_6$* ***3.116***: The peptide in the solution above is diluted with 700 µl $H_2O$+ 0.1% TFA and purified on a semi-preparative HPLC using a gradient of 03-60% B in 30 min. Combined fractions are united and lyophilized to yield the pure MUC1(1-80)-His$_6$ **5** as a white solid.

**3.116** H-MUC1(1-80)-His$_6$: 1.5 mg, 8316 gr/mol, 0.19 µmol, 45%.

ESI-MS (positive mode) m/z = 1189.5 ($C_{356}H_{552}N_{119}O_{114}$ $[M+7H]^{7+}$, calcd. 1189.0), 1040.5 ($C_{356}H_{553}N_{119}O_{114}$ $[M+8H]^{8+}$, calcd. 1040.5, 925.0 ($C_{356}H_{554}N_{119}O_{114}$ $[M+9H]^{9+}$, calcd.925.0), 832.6 ($C_{356}H_{555}N_{119}O_{114}$ $[M+10H]^{10+}$, calcd. 832.6), 757.0 ($C_{356}H_{556}N_{119}O_{114}$ $[M+11H]^{11+}$, calcd. 757.0), 694.0 (C $C_{356}H_{557}N_{119}O_{114}$ $[M+12H]^{12+}$, calcd. 694.0), 640.0 ($C_{356}H_{558}N_{119}O_{114}$ $[M+13H]^{13+}$, calcd. 640.7), 595.0 ($C_{356}H_{559}N_{119}O_{114}$ $[M+14H]^{14+}$, calcd. 595.0).

# References


1. Boyle, J., *Lehninger principles of biochemistry (4th ed.): Nelson, D., and Cox, M.* Biochemistry and Molecular Biology Education, 2005. **33**(1): p. 74-75.
2. Pentelute, B. 2008, Chicago: Chicago.
3. Gamblin, D.P., E.M. Scanlan, and B.G. Davis, *Glycoprotein Synthesis: An Update.* Chemical Reviews, 2009. **109**(1): p. 131-163.
4. Malins, L.R., K.M. Cergol, and R.J. Payne, *Chemoselective sulfenylation and peptide ligation at tryptophan.* Chemical Science, 2014. **5**(1): p. 260-266.
5. McGinty, R.K., et al., *Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation.* Nature, 2008. **453**(7196): p. 812-6.
6. Bondalapati, S., et al., *Chemical synthesis of phosphorylated ubiquitin and diubiquitin exposes positional sensitivities of e1-e2 enzymes and deubiquitinases.* Chemistry, 2015. **21**(20): p. 7360-4.
7. Okamoto, R., et al., *Total Chemical Synthesis and Biological Activities of Glycosylated and Non-Glycosylated Forms of the Chemokines CCL1 and Ser-CCL1.* Angewandte Chemie-International Edition, 2014. **53**(20): p. 5188-5193.
8. Bock, V.D., et al., *Click chemistry as a route to cyclic tetrapeptide analogues: synthesis of cyclo-[Pro-Val-psi(triazole)-Pro-Tyr].* Org Lett, 2006. **8**(5): p. 919-22.
9. Dirksen, A., T.M. Hackeng, and P.E. Dawson, *Nucleophilic catalysis of oxime ligation.* Angew Chem Int Ed Engl, 2006. **45**(45): p. 7581-4.
10. Saxon, E. and C.R. Bertozzi, *Cell surface engineering by a modified Staudinger reaction.* Science, 2000. **287**(5460): p. 2007-2010.
11. Soellner, M.B., A. Tam, and R.T. Raines, *Staudinger ligation of peptides at non-glycyl residues.* Journal of Organic Chemistry, 2006. **71**(26): p. 9824-9830.
12. Bode, J.W., R.M. Fox, and K.D. Baucom, *Chemoselective amide ligations by decarboxylative condensations of N-alkylhydroxylamines and alpha-ketoacids.* Angewandte Chemie-International Edition, 2006. **45**(8): p. 1248-1252.
13. Merrifield, R.B., *Solid Phase Peptide Synthesis .1. Synthesis of a Tetrapeptide.* Journal of the American Chemical Society, 1963. **85**(14): p. 2149-&.
14. Kochendoerfer, G.G. and S.B.H. Kent, *Chemical protein synthesis.* Current Opinion in Chemical Biology, 1999. **3**(6): p. 665-671.
15. Ball, H.L. and P. Mascagni, *Chemical synthesis and purification of proteins: A methodology.* International Journal of Peptide and Protein Research, 1996. **48**(1): p. 31-47.
16. El Oualid, F., et al., *Chemical Synthesis of Ubiquitin, Ubiquitin-Based Probes, and Diubiquitin.* Angewandte Chemie International Edition, 2010. **49**(52): p. 10149-10153.
17. Schnolzer, M. and S.B.H. Kent, *Constructing Proteins by Dovetailing Unprotected Synthetic Peptides - Backbone-Engineered Hiv Protease.* Science, 1992. **256**(5054): p. 221-225.
18. Englebretsen, D.R., et al., *A Novel Thioether Linker - Chemical Synthesis of a Hiv-1 Protease Analog by Thioether Ligation.* Tetrahedron Letters, 1995. **36**(48): p. 8871-8874.
19. Liu, C.F. and J.P. Tam, *Chemical Ligation Approach to Form a Peptide-Bond between Unprotected Peptide Segments - Concept and Model Study.* Journal of the American Chemical Society, 1994. **116**(10): p. 4149-4153.

20. Rose, K., *Facile Synthesis of Homogeneous Artificial Proteins.* Journal of the American Chemical Society, 1994. **116**(1): p. 30-33.
21. Fotouhi, N., N.G. Galakatos, and D.S. Kemp, *Peptide-Synthesis by Prior Thiol Capture .6. Rates of the Disulfide Bond Forming Capture Reaction and Demonstration of the Overall Strategy by Synthesis of the C-Terminal 29-Peptide Sequence of Bpti.* Journal of Organic Chemistry, 1989. **54**(12): p. 2803-2817.
22. Wieland, T., et al., *Polypeptide syntheses. VIII. Formation of sulfur containing peptides by the intramolecular migration of aminoacyl groups.* Justus Liebigs Ann. Chem., 1953. **583**: p. 129-49.
23. Dawson, P.E., et al., *Synthesis of Proteins by Native Chemical Ligation.* Science, 1994. **266**(5186): p. 776-779.
24. Schmalisch, J. and O. Seitz, *Acceleration of thiol additive-free native chemical ligation by intramolecular S -> S acyl transfer.* Chemical Communications, 2015. **51**(35): p. 7554-7557.
25. Haase, C. and O. Seitz, *Internal Cysteine Accelerates Thioester-Based Peptide Ligation.* European Journal of Organic Chemistry, 2009(13): p. 2096-2101.
26. McCaldon, P. and P. Argos, *Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences.* Proteins: Structure, Function, and Bioinformatics, 1988. **4**(2): p. 99-122.
27. Beligere, G.S. and P.E. Dawson, *Conformationally assisted protein ligation using C-terminal thioester peptides.* Journal of the American Chemical Society, 1999. **121**(26): p. 6332-6333.
28. Tam, J.P. and Q.T. Yu, *Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones.* Biopolymers, 1998. **46**(5): p. 319-327.
29. Bochar, D.A., et al., *Aminoethylcysteine can replace the function of the essential active site lysine of Pseudomonas mevalonii 3-hydroxy-3-methylglutaryl coenzyme A reductase.* Biochemistry, 1999. **38**(28): p. 8879-8883.
30. Kochendoerfer, G.G., et al., *Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein.* Science, 2003. **299**(5608): p. 884-887.
31. Okamoto, R. and Y. Kajihara, *Uncovering a latent ligation site for glycopeptide synthesis.* Angewandte Chemie-International Edition, 2008. **47**(29): p. 5402-5406.
32. Yan, L.Z. and P.E. Dawson, *Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization.* Journal of the American Chemical Society, 2001. **123**(4): p. 526-533.
33. Wan, Q. and S.J. Danishefsky, *Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides.* Angewandte Chemie-International Edition, 2007. **46**(48): p. 9248-9252.
34. Crich, D. and A. Banerjee, *Native chemical ligation at phenylalanine.* Journal of the American Chemical Society, 2007. **129**(33): p. 10064-+.
35. Malins, L.R., et al., *Synthesis of β-Thiol Phenylalanine for Applications in One-Pot Ligation–Desulfurization Chemistry.* Organic Letters, 2015. **17**(9): p. 2070-2073.
36. Haase, C., H. Rohde, and O. Seitz, *Native chemical ligation at valine.* Angewandte Chemie-International Edition, 2008. **47**(36): p. 6807-6810.
37. Chen, J., et al., *A program for ligation at threonine sites: application to the controlled total synthesis of glycopeptides.* Tetrahedron, 2010. **66**(13): p. 2277-2283.

38. Kumar, K.S.A., et al., *Highly Efficient and Chemoselective Peptide Ubiquitylation.* Angewandte Chemie-International Edition, 2009. **48**(43): p. 8090-8094.
39. Yang, R.L., et al., *Dual Native Chemical Ligation at Lysine.* Journal of the American Chemical Society, 2009. **131**(38): p. 13592-+.
40. Harpaz, Z., et al., *Protein Synthesis Assisted by Native Chemical Ligation at Leucine.* Chembiochem, 2010. **11**(9): p. 1232-1235.
41. Tan, Z.P., S.Y. Shang, and S.J. Danishefsky, *Insights into the Finer Issues of Native Chemical Ligation: An Approach to Cascade Ligations.* Angewandte Chemie-International Edition, 2010. **49**(49): p. 9500-9503.
42. Shang, S.Y., et al., *An Advance in Proline Ligation.* Journal of the American Chemical Society, 2011. **133**(28): p. 10784-10786.
43. Ding, H., et al., *Dual Kinetically Controlled Native Chemical Ligation Using a Combination of Sulfanylproline and Sulfanylethylanilide Peptide.* Organic Letters, 2011. **13**(20): p. 5588-5591.
44. Siman, P., S.V. Karthikeyan, and A. Brik, *Native Chemical Ligation at Glutamine.* Organic Letters, 2012. **14**(6): p. 1520-1523.
45. Cergol, K.M., et al., *One-Pot Peptide Ligation-Desulfurization at Glutamate.* Organic Letters, 2014. **16**(1): p. 290-293.
46. Thompson, R.E., et al., *Chemoselective Peptide Ligation-Desulfurization at Aspartate.* Angewandte Chemie-International Edition, 2013. **52**(37): p. 9723-9727.
47. Malins, L.R., K.M. Cergol, and R.J. Payne, *Peptide Ligation-Desulfurization Chemistry at Arginine.* Chembiochem, 2013. **14**(5): p. 559-563.
48. Moyal, T., et al., *Highly efficient one-pot ligation and desulfurization.* Chemical Science, 2013. **4**(6): p. 2496-2501.
49. Thompson, R.E., et al., *Trifluoroethanethiol: An Additive for Efficient One-Pot Peptide Ligation-Desulfurization Chemistry.* Journal of the American Chemical Society, 2014. **136**(23): p. 8161-8164.
50. Wang, P., et al., *At Last: Erythropoietin as a Single Glycoform.* Angewandte Chemie-International Edition, 2012. **51**(46): p. 11576-11584.
51. Kumar, K.S.A., et al., *Total Chemical Synthesis of a 304 Amino Acid K48-Linked Tetraubiquitin Protein.* Angewandte Chemie-International Edition, 2011. **50**(27): p. 6137-6141.
52. Brik, A., et al., *Sugar-assisted glycopeptide ligation.* Journal of the American Chemical Society, 2006. **128**(17): p. 5626-5627.
53. Dwek, R.A., *Glycobiology: Toward understanding the function of sugars.* Chemical Reviews, 1996. **96**(2): p. 683-720.
54. Brik, A., S. Ficht, and C.H. Wong, *Strategies for the preparation of homogenous glycoproteins.* Current Opinion in Chemical Biology, 2006. **10**(6): p. 638-644.
55. Brik, A., et al., *Sugar-assisted ligation of N-linked glycopeptides with broad sequence tolerance at the ligation junction.* Journal of the American Chemical Society, 2006. **128**(46): p. 15026-15033.
56. Ficht, S., et al., *Second-generation sugar-assisted ligation: A method for the synthesis of cysteine-containing glycopeptides.* Angewandte Chemie-International Edition, 2007. **46**(31): p. 5975-5979.
57. Payne, R.J., et al., *Extended sugar-assisted glycopeptide ligations: Development, scope, and applications.* Journal of the American Chemical Society, 2007. **129**(44): p. 13527-13536.

58. Lutsky, M.Y., N. Nepomniaschiy, and A. Brik, *Peptide ligation via side-chain auxiliary.* Chemical Communications, 2008(10): p. 1229-1231.
59. Kumar, K.S.A., et al., *Side-chain assisted ligation in protein synthesis.* Bioorganic & Medicinal Chemistry Letters, 2009. **19**(14): p. 3870-3874.
60. Canne, L.E., S.J. Bark, and S.B.H. Kent, *Extending the applicability of native chemical ligation.* Journal of the American Chemical Society, 1996. **118**(25): p. 5891-5896.
61. Cardona, V.M.F., O. Hartley, and P. Botti, *Synthesis of cyclic peptides from unprotected precursors using removable N-alpha-(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary.* Journal of Peptide Research, 2003. **61**(3): p. 152-157.
62. Botti, P., M.R. Carrasco, and S.B.H. Kent, *Native chemical ligation using removable N-alpha-(1-phenyl-2-mercaptoethyl) auxiliaries.* Tetrahedron Letters, 2001. **42**(10): p. 1831-1833.
63. Low, D.W., et al., *Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(12): p. 6554-6559.
64. Marinzi, C., et al., *A new scaffold for amide ligation.* Bioorganic & Medicinal Chemistry, 2001. **9**(9): p. 2323-2328.
65. Kawakami, T., K. Akaji, and S. Aimoto, *Peptide bond formation mediated by 4,5-dimethoxy-2-mercaptobenzylamine after periodate oxidation of the N-terminal serine residue.* Organic Letters, 2001. **3**(9): p. 1403-1405.
66. Offer, J. and P.E. Dawson, *N-alpha-2-mercaptobenzylamine-assisted chemical ligation.* Organic Letters, 2000. **2**(1): p. 23-26.
67. Offer, J., C.N.C. Boddy, and P.E. Dawson, *Extending synthetic access to proteins with a removable acyl transfer auxiliary.* Journal of the American Chemical Society, 2002. **124**(17): p. 4642-4646.
68. Kawakami, T. and S. Aimoto, *A photoremovable ligation auxiliary for use in polypeptide synthesis.* Tetrahedron Letters, 2003. **44**(32): p. 6059-6061.
69. Marinzi, C., et al., *An o-nitrobenzyl scaffold for peptide ligation: synthesis and applications.* Bioorganic & Medicinal Chemistry, 2004. **12**(10): p. 2749-2757.
70. Nadler, C., et al., *A Photocleavable Auxiliary for Extended Native Chemical Ligation.* European Journal of Organic Chemistry, 2015(14): p. 3095-3102.
71. Staudinger, H. and J. Meyer, *Über neue organische Phosphorverbindungen III. Phosphinmethylenderivate und Phosphinimine.* Helvetica Chimica Acta, 1919. **2**(1): p. 635-646.
72. Tam, A., M.B. Soellner, and R.T. Raines, *Water-soluble phosphinothiols for traceless Staudinger ligation and integration with expressed protein ligation.* Journal of the American Chemical Society, 2007. **129**(37): p. 11421-11430.
73. Nilsson, B.L., et al., *Protein assembly by orthogonal chemical ligation methods.* J Am Chem Soc, 2003. **125**(18): p. 5268-9.
74. David, O., et al., *Intramolecular Staudinger ligation: a powerful ring-closure method to form medium-sized lactams.* Angew Chem Int Ed Engl, 2003. **42**(36): p. 4373-5.
75. Pattabiraman, V.R., A.O. Ogunkoya, and J.W. Bode, *Chemical protein synthesis by chemoselective alpha-ketoacid-hydroxylamine (KAHA) ligations with 5-oxaproline.* Angew Chem Int Ed Engl, 2012. **51**(21): p. 5114-8.
76. Pusterla, I. and J.W. Bode, *An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations.* Nat Chem, 2015. **7**(8): p. 668-72.

77. Wucherpfennig, T.G., et al., *Traceless preparation of C-terminal alpha-ketoacids for chemical protein synthesis by alpha-ketoacid-hydroxylamine ligation: synthesis of SUMO2/3.* Angew Chem Int Ed Engl, 2014. **53**(45): p. 12248-52.
78. Ogunkoya, A.O., V.R. Pattabiraman, and J.W. Bode, *Sequential alpha-ketoacid-hydroxylamine (KAHA) ligations: synthesis of C-terminal variants of the modifier protein UFM1.* Angew Chem Int Ed Engl, 2012. **51**(38): p. 9693-7.
79. He, C., et al., *Chemical Synthesis of the 20 kDa Heme Protein Nitrophorin 4 by alpha-Ketoacid-Hydroxylamine (KAHA) Ligation.* Angew Chem Int Ed Engl, 2015. **54**(44): p. 12996-3001.
80. Bang, D. and S.B.H. Kent, *A one-pot total synthesis of crambin.* Angewandte Chemie-International Edition, 2004. **43**(19): p. 2534-2538.
81. Boerema, D.J., V.A. Tereshko, and S.B.H. Kent, *Total synthesis by modern chemical Ligation methods and high resolution (1.1 angstrom) X-ray structure of ribonuclease A.* Biopolymers, 2008. **90**(3): p. 278-286.
82. Sohma, Y., et al., *Comparative properties of insulin-like growth factor 1 (IGF-1) and [Gly7D-Ala]IGF-1 prepared by total chemical synthesis.* Angewandte Chemie-International Edition, 2008. **47**(6): p. 1102-1106.
83. Mandal, K. and S.B.H. Kent, *Total Chemical Synthesis of Biologically Active Vascular Endothelial Growth Factor.* Angewandte Chemie-International Edition, 2011. **50**(35): p. 8029-8033.
84. Mhidia, R., et al., *Synthesis of peptide-protein conjugates using N-succinimidyl carbamate chemistry.* Bioconjug Chem, 2010. **21**(2): p. 219-28.
85. Ollivier, N., et al., *Bis(2-sulfanylethyl)amino native peptide ligation.* Org Lett, 2010. **12**(22): p. 5238-41.
86. Ollivier, N., et al., *A One-Pot Three-Segment Ligation Strategy for Protein Chemical Synthesis.* Angewandte Chemie-International Edition, 2012. **51**(1): p. 209-213.
87. Bang, D., B.L. Pentelute, and S.B.H. Kent, *Kinetically controlled ligation for the convergent chemical synthesis of proteins.* Angewandte Chemie-International Edition, 2006. **45**(24): p. 3985-3988.
88. Kawakami, T. and S. Aimoto, *Sequential peptide ligation by using a controlled cysteinyl prolyl ester (CPE) autoactivating unit.* Tetrahedron Letters, 2007. **48**(11): p. 1903-1905.
89. Erlich, L.A., et al., *N-Methylcysteine-mediated total chemical synthesis of ubiquitin thioester.* Organic & Biomolecular Chemistry, 2010. **8**(10): p. 2392-2396.
90. Zheng, J.S., et al., *Chemical Protein Synthesis by Kinetically Controlled Ligation of Peptide O-Esters.* Chembiochem, 2010. **11**(4): p. 511-515.
91. Sato, K., et al., *N-Sulfanylethylanilide Peptide as a Crypto-Thioester Peptide.* Chembiochem, 2011. **12**(12): p. 1840-1844.
92. Canne, L.E., et al., *Chemical protein synthesis by solid phase ligation of unprotected peptide segments.* Journal of the American Chemical Society, 1999. **121**(38): p. 8720-8727.
93. Raibaut, L., et al., *Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments.* Chemical Science, 2013. **4**(10): p. 4061-4066.
94. Jbara, M., M. Seenaiah, and A. Brik, *Solid phase chemical ligation employing a rink amide linker for the synthesis of histone H2B protein.* Chemical Communications, 2014. **50**(83): p. 12534-12537.

95. Loibl, S.H., Ziv; Seitz, Oliver, *A Type of Auxiliary for Native Chemical Peptide Ligation beyond Cysteine and Glycine Junctions.* Angewandte Chemie International Edition, 2015. **54**: p. 1-6.
96. Christian, H., *Neue Auxiliare für die Peptidfragmentverknüpfung*, 2009, Humboldt Universität zu Berlin.
97. Choi, H.Y. and D.Y. Chi, *Nonselective bromination-selective debromination strategy: Selective bromination of unsymmetrical ketones on singly activated carbon against doubly activated carbon.* Organic Letters, 2003. **5**(4): p. 411-414.
98. Dijkstra G., K.W.M., Kellogg R. M. , *An Assessment of the Cause of the "Cesium Effect".* J. Org. Chem, 1987. **52**: p. 4230-4234.
99. Martin, S.F., M.P. Dwyer, and C.L. Lynch, *Application of AlMe3-mediated amidation reactions to solution phase peptide synthesis.* Tetrahedron Letters, 1998. **39**(12): p. 1517-1520.
100. Dess, D.B. and J.C. Martin, *Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones.* The Journal of Organic Chemistry, 1983. **48**(22): p. 4155-4156.
101. Biosystems, A., *Introduction to Cleavage Techniques*. 1995, Applied Biosystems.
102. Offer, J., *Native Chemical Ligation with N-alpha Acyl Transfer Auxiliaries.* Biopolymers, 2010. **94**(4): p. 530-541.
103. Schittek, B., et al., *Dermcidin: a novel human antibiotic peptide secreted by sweat glands.* Nature Immunology, 2001. **2**(12): p. 1133-1137.
104. Porter, D., et al., *A neural survival factor is a candidate oncogene in breast cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(19): p. 10931-10936.
105. Lowrie, A.G., et al., *Dermcidin expression in hepatic cells improves survival without N-glycosylation, but requires asparagine residues.* British Journal of Cancer, 2006. **94**(11): p. 1663-1671.
106. Minami, Y., et al., *Immunohistochemical staining of cutaneous tumours with G-81, a monoclonal antibody to dermcidin.* British Journal of Dermatology, 2004. **151**(1): p. 165-169.
107. Miniami, Y., et al., *Cutaneous mixed tumors: an immunohistochemical study using two antibodies, G-81 and C8/144B.* Journal of Dermatological Science, 2004. **36**(3): p. 180-182.
108. Rieg, S., et al., *Dermcidin is constitutively produced by eccrine sweat glands and is not induced in epidermal cells under inflammatory skin conditions.* British Journal of Dermatology, 2004. **151**(3): p. 534-539.
109. Stewart, G.D., et al., *Dermcidin expression confers a survival advantage in prostate cancer cells subjected to oxidative stress or hypoxia.* Prostate, 2007. **67**(12): p. 1308-1317.
110. Stewart, G.D., et al., *Variation in dermcidin expression in a range of primary human tumours and in hypoxic/oxidatively stressed human cell lines.* British Journal of Cancer, 2008. **99**(1): p. 126-132.
111. Stewart, G.D., et al., *The dermcidin gene in cancer: role in cachexia, carcinogenesis and tumour cell survival.* Current Opinion in Clinical Nutrition and Metabolic Care, 2008. **11**(3): p. 208-213.
112. Cunningham, T.J., et al., *Identification of a survival-promoting peptide in medium conditioned by oxidatively stressed cell lines of nervous system origin.* Journal of Neuroscience, 1998. **18**(18): p. 7047-7060.

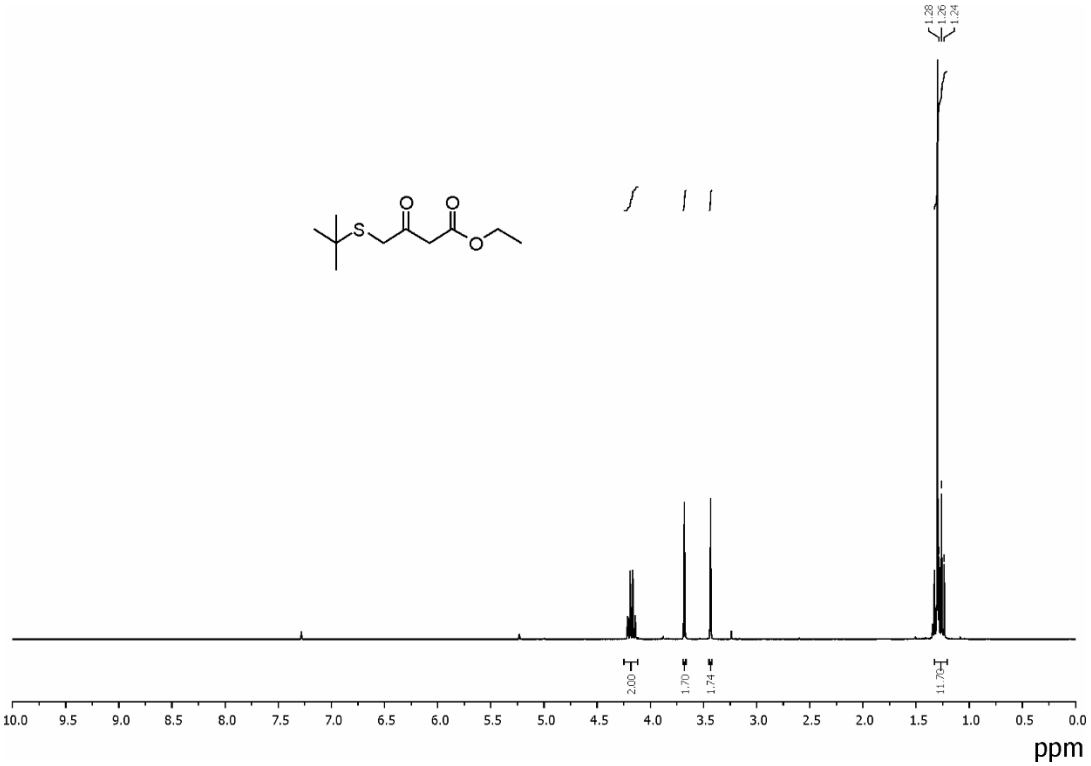
113. Cunningham, T.J., et al., *Identification of the human cDNA for new survival/evasion peptide (DSEP): Studies in vitro and in vivo of overexpression by neural cells.* Experimental Neurology, 2002. **177**(1): p. 32-39.
114. Rieg, S., et al., *Generation of multiple stable dermcidin-derived antimicrobial peptides in sweat of different body sites.* Journal of Investigative Dermatology, 2006. **126**(2): p. 354-365.
115. Steffen, H., et al., *Naturally processed dermcidin-derived peptides do not permeabilize bacterial membranes and kill microorganisms irrespective of their charge.* Antimicrobial Agents and Chemotherapy, 2006. **50**(8): p. 2608-2620.
116. Todorov, P.T., W.N. Field, and M.J. Tisdale, *Role of a proteolysis inducing factor (PIF) in cachexia induced by a human melanoma (G361).* British Journal of Cancer, 1999. **80**(11): p. 1734-1737.
117. Wigmore, S.J., et al., *Characteristics of patients with pancreatic cancer expressing a novel cancer cachectic factor.* British Journal of Surgery, 2000. **87**(1): p. 53-58.
118. Lee Motoyama, J.P., et al., *Identification of dermcidin in human gestational tissue and characterization of its proteolytic activity.* Biochem Biophys Res Commun, 2007. **357**(4): p. 828-33.
119. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.
120. *AMPer: A Database and an Automated Discovery Tool for Gene-Coded Antimicrobial Peptides*. Available from: http://marray.cmdr.ubc.ca/cgi-bin/amp.pl?peptide=1&type=MATURE.
121. Ingenito, R., et al., *Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc/t-Bu Chemistry.* Journal of the American Chemical Society, 1999. **121**(49): p. 11369-11374.
122. Mehla, K. and P.K. Singh, *MUC1: a novel metabolic master regulator.* Biochim Biophys Acta, 2014. **1845**(2): p. 126-35.
123. Yolken, R.H., et al., *Human milk mucin inhibits rotavirus replication and prevents experimental gastroenteritis.* J Clin Invest, 1992. **90**(5): p. 1984-91.
124. Kufe, D.W., *Mucins in cancer: function, prognosis and therapy.* Nat Rev Cancer, 2009. **9**(12): p. 874-85.
125. Hochuli, E., et al., *Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent.* Bio-Technology, 1988. **6**(11): p. 1321-1325.
126. Hengen, P.N., *Methods and Reagents - Purification of His-Tag Fusion Proteins from Escherichia-Coli.* Trends in Biochemical Sciences, 1995. **20**(7): p. 285-286.
127. Bang, D. and S.B.H. Kent, *HiS(6) tag-assisted chemical protein synthesis.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(14): p. 5014-5019.
128. Blanco-Canosa, J.B. and P.E. Dawson, *An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation.* Angewandte Chemie-International Edition, 2008. **47**(36): p. 6851-6855.
129. Durek, T., V.Y. Torbeev, and S.B.H. Kent, *Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(12): p. 4846-4851.
130. Fang, G.M., et al., *Protein Chemical Synthesis by Ligation of Peptide Hydrazides.* Angewandte Chemie-International Edition, 2011. **50**(33): p. 7645-7649.

131. Fang, G.M., J.X. Wang, and L. Liu, *Convergent chemical synthesis of proteins by ligation of peptide hydrazides.* Angew Chem Int Ed Engl, 2012. **51**(41): p. 10347-50.
132. Johnson, E.C.B. and S.B.H. Kent, *Insights into the mechanism and catalysis of the native chemical ligation reaction.* Journal of the American Chemical Society, 2006. **128**(20): p. 6640-6646.
133. Romovacek, H., et al., *Studies on Polypeptides .54. Synthesis of a Peptide Corresponding to Positions 24-104 of the Peptide-Chain of Ribonuclease-T1.* Journal of the American Chemical Society, 1979. **101**(20): p. 6081-6091.
134. Felix, A.M. and R.B. Merrifield, *Azide solid phase peptide synthesis.* J Am Chem Soc, 1970. **92**(5): p. 1385-91.
135. Yanaihara, N., et al., *Studies on polypeptides. XLII. Synthesis and characterization of seven fragments spanning the entire sequence of ribonuclease T1.* J Am Chem Soc, 1969. **91**(8): p. 2184-5.
136. Chundak, S.J., V.M. Leovac, and L.J. Bjelica, *Coordination-Compounds of Co(Ii), Ni(Ii) and Cu(Ii) with Capric Acid Hydrazide.* Monatshefte Fur Chemie, 1987. **118**(8-9): p. 923-930.
137. Hojo, H., et al., *N-alkyl cysteine-assisted thioesterification of peptides (vol 48, pg 25, 2007).* Tetrahedron Letters, 2007. **48**(7): p. 1299-1299.
138. Hou, W., et al., *Peptidyl N,N-Bis(2-mercaptoethyl)-amides as Thioester Precursors for Native Chemical Ligation.* Organic Letters, 2011. **13**(3): p. 386-389.
139. Tsuda, S., et al., *N -> S Acyl-Transfer-Mediated Synthesis of Peptide Thioesters Using Anilide Derivatives.* Organic Letters, 2009. **11**(4): p. 823-826.
140. Mende, F., M. Beisswenger, and O. Seitz, *Automated Fmoc-Based Solid-Phase Synthesis of Peptide Thioesters with Self-Purification Effect and Application in the Construction of Immobilized SH3 Domains.* Journal of the American Chemical Society, 2010. **132**(32): p. 11110-11118.
141. Schnolzer, M., et al., *In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences.* Int J Pept Protein Res, 1992. **40**(3-4): p. 180-93.

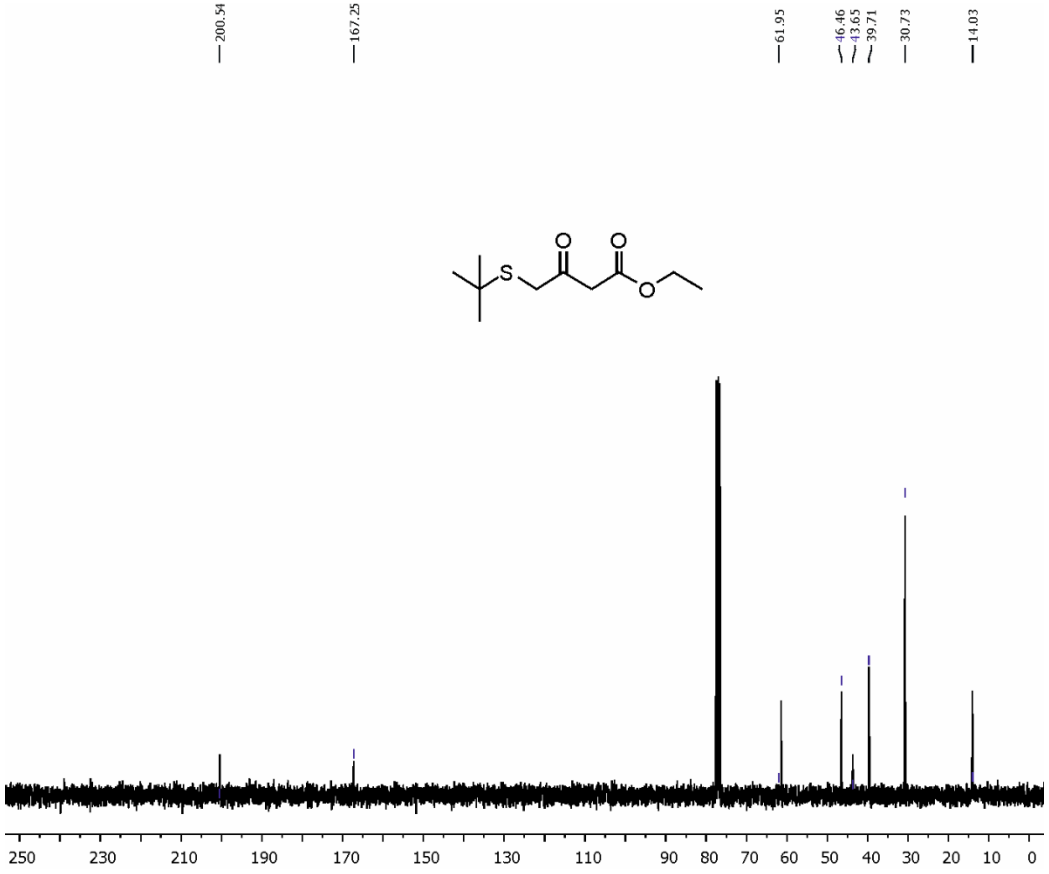
# Appendix A: Spectrometry Data

## A.1 NMR Spectra

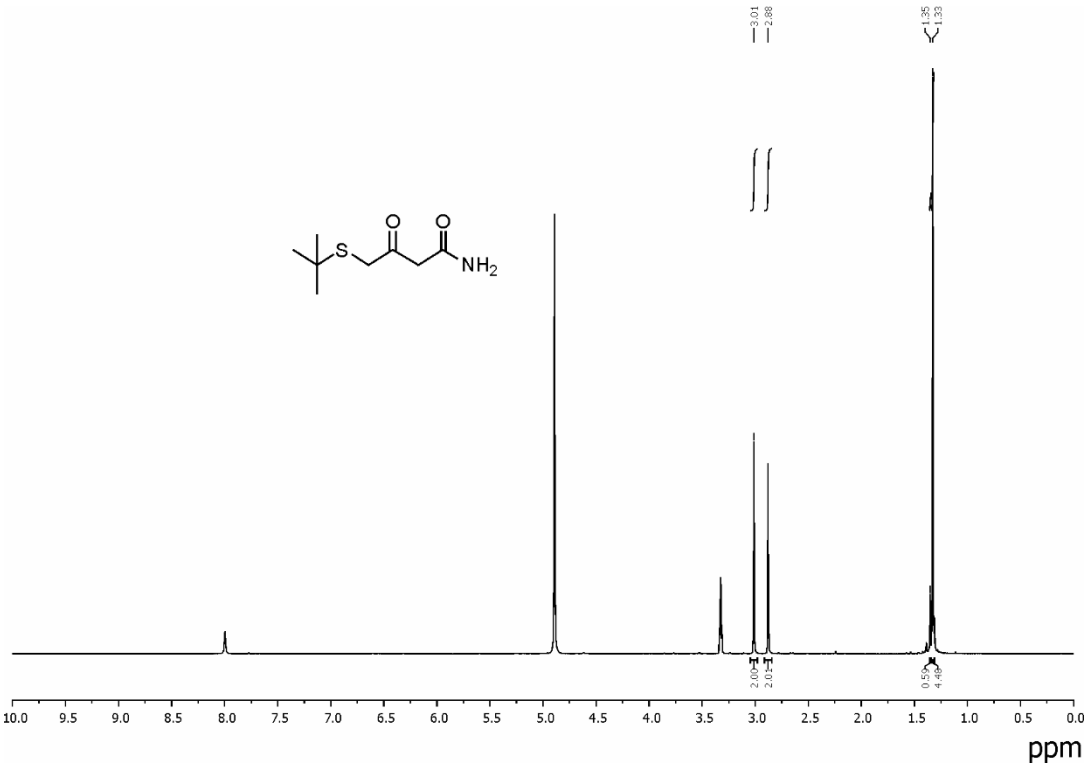
$^{1}$H NMR **3.021**



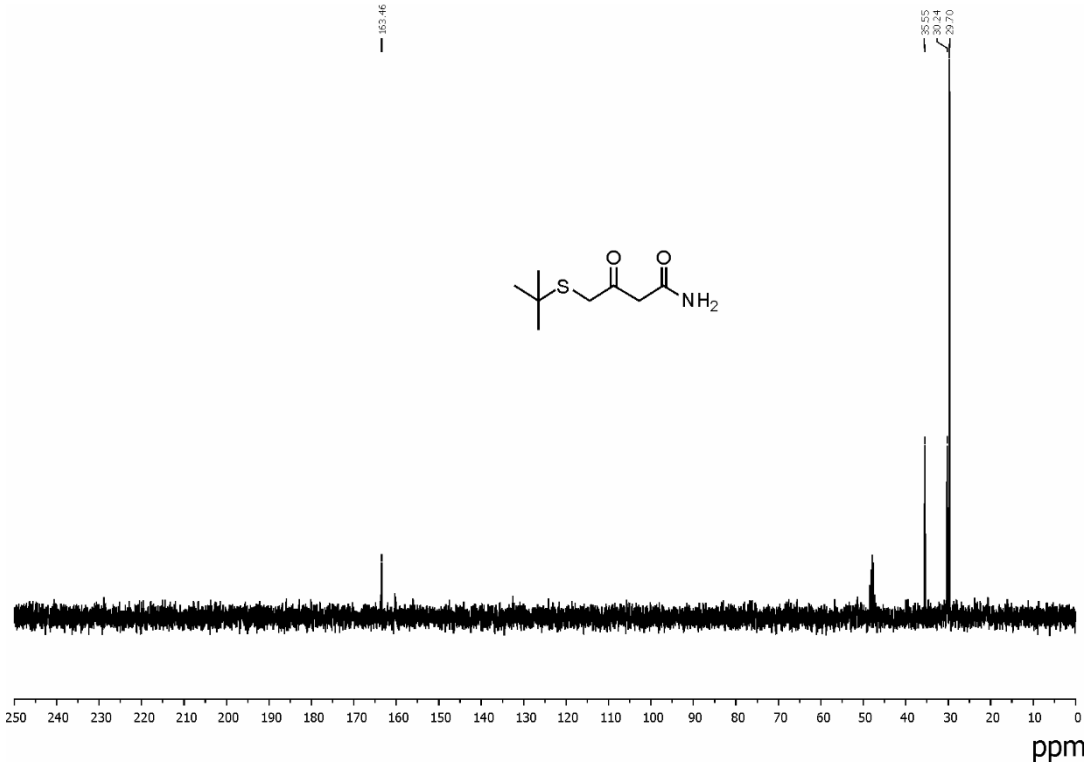

$^{13}$C NMR **3.021**



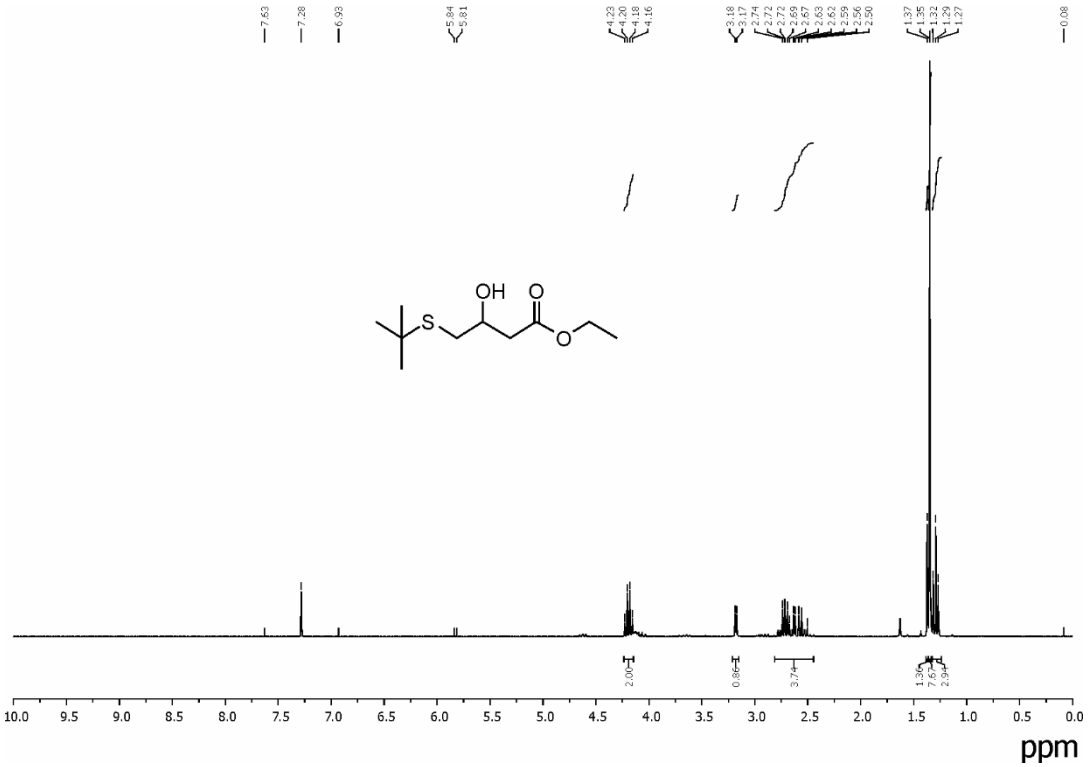
$^{1}$H NMR **3.024**



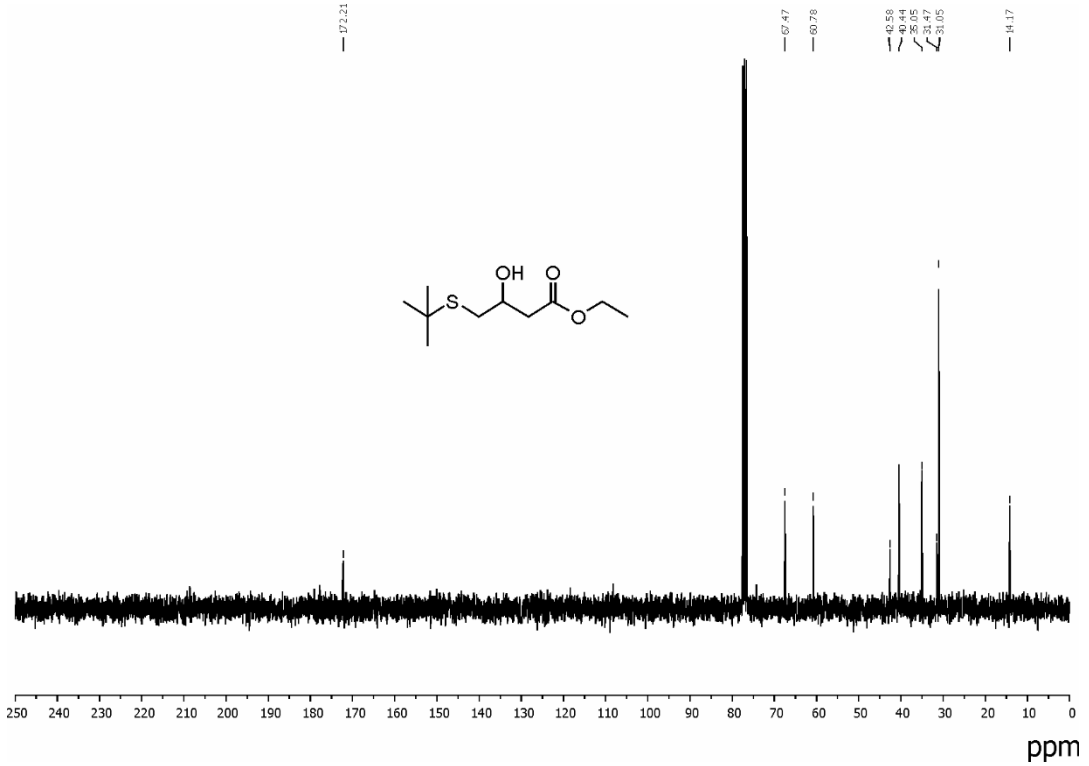

$^{13}$C NMR **3.024**



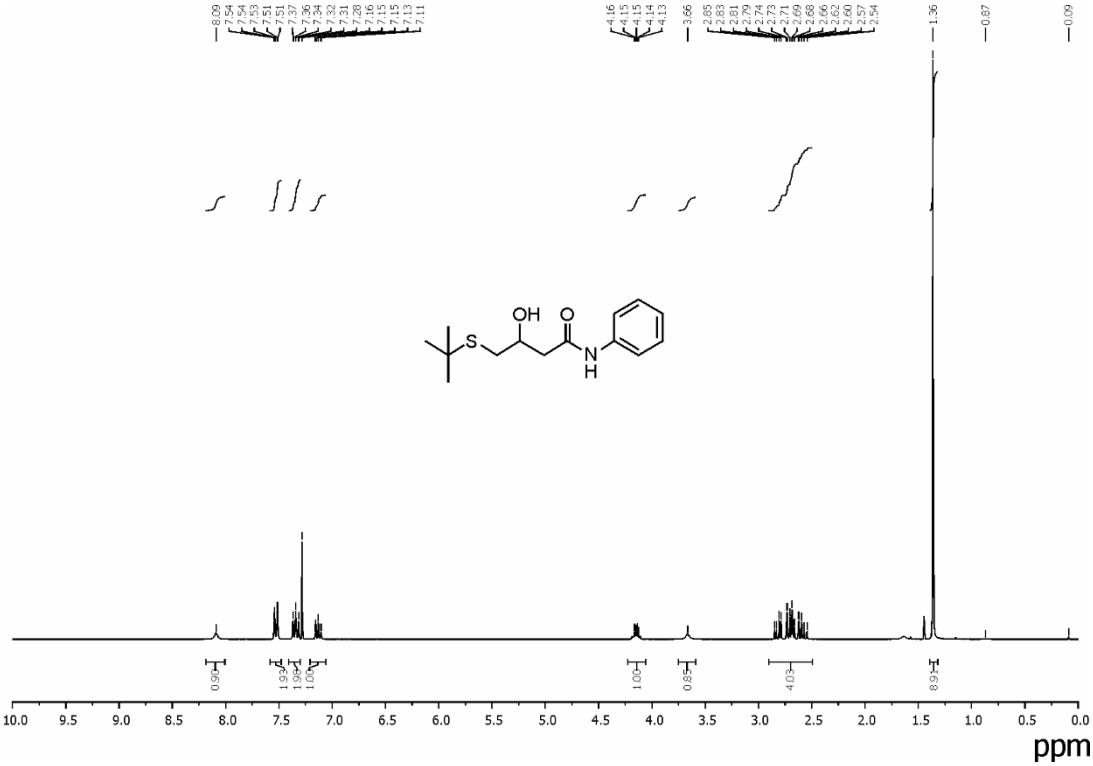
$^{1}$H NMR **3.026**



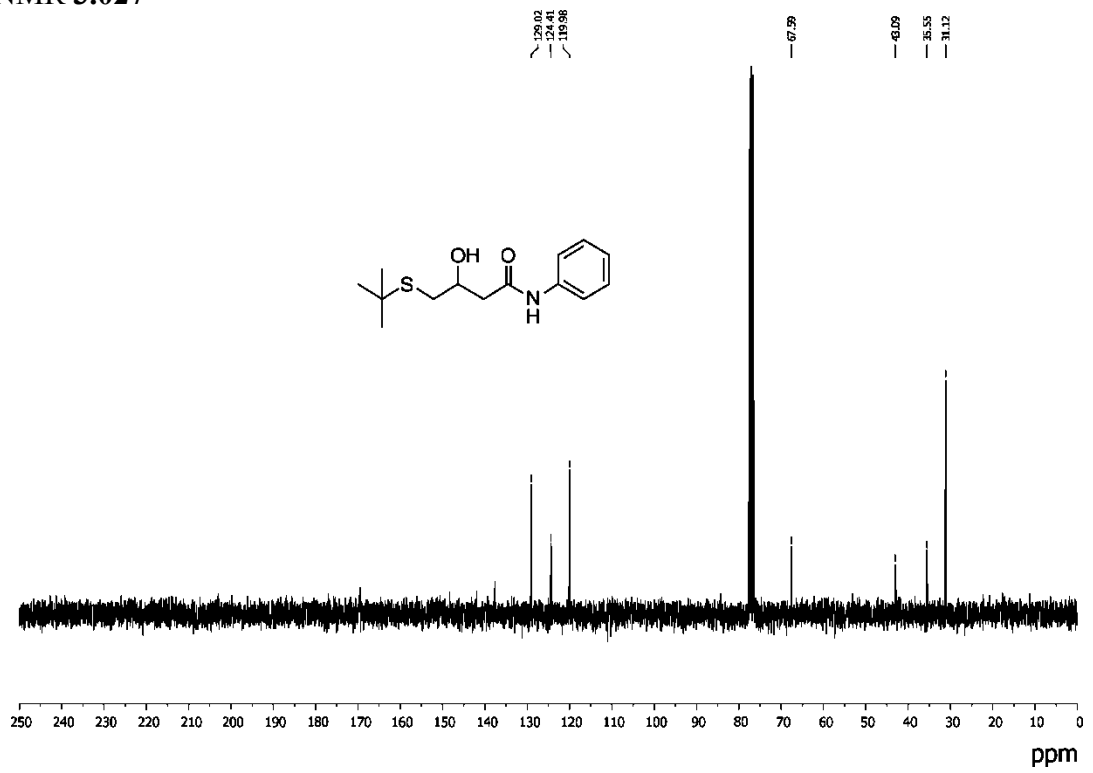

$^{13}$C NMR **3.026**



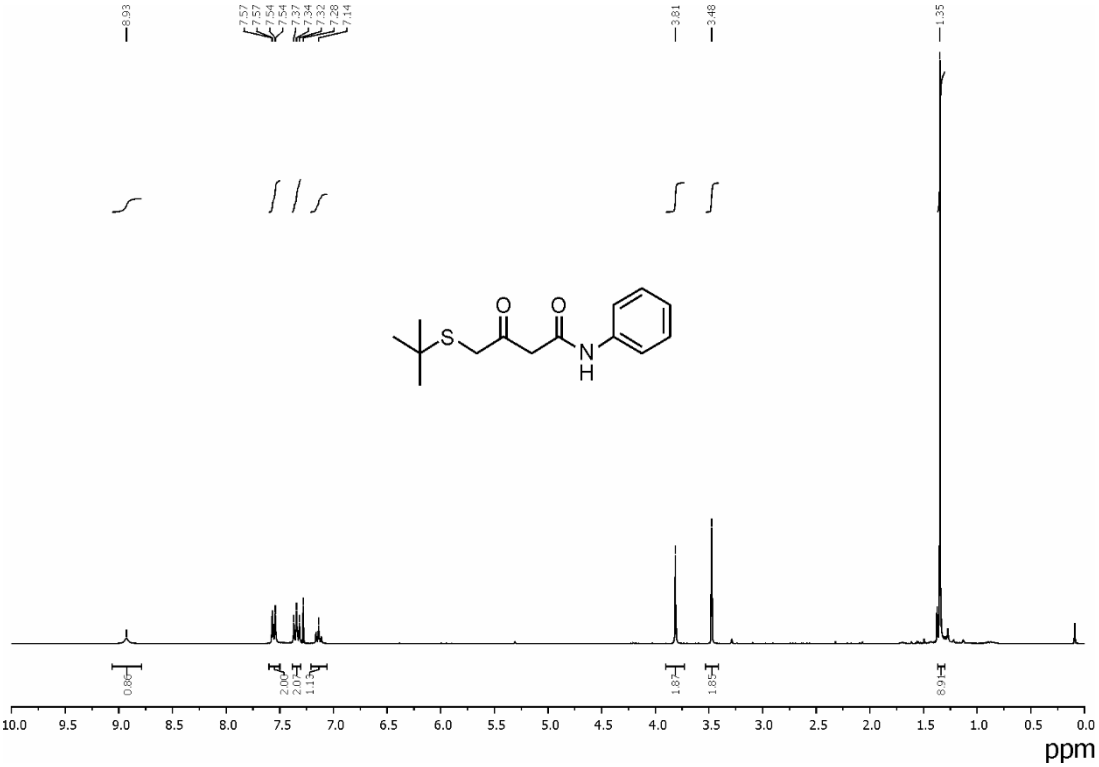
$^{1}$H NMR **3.027**



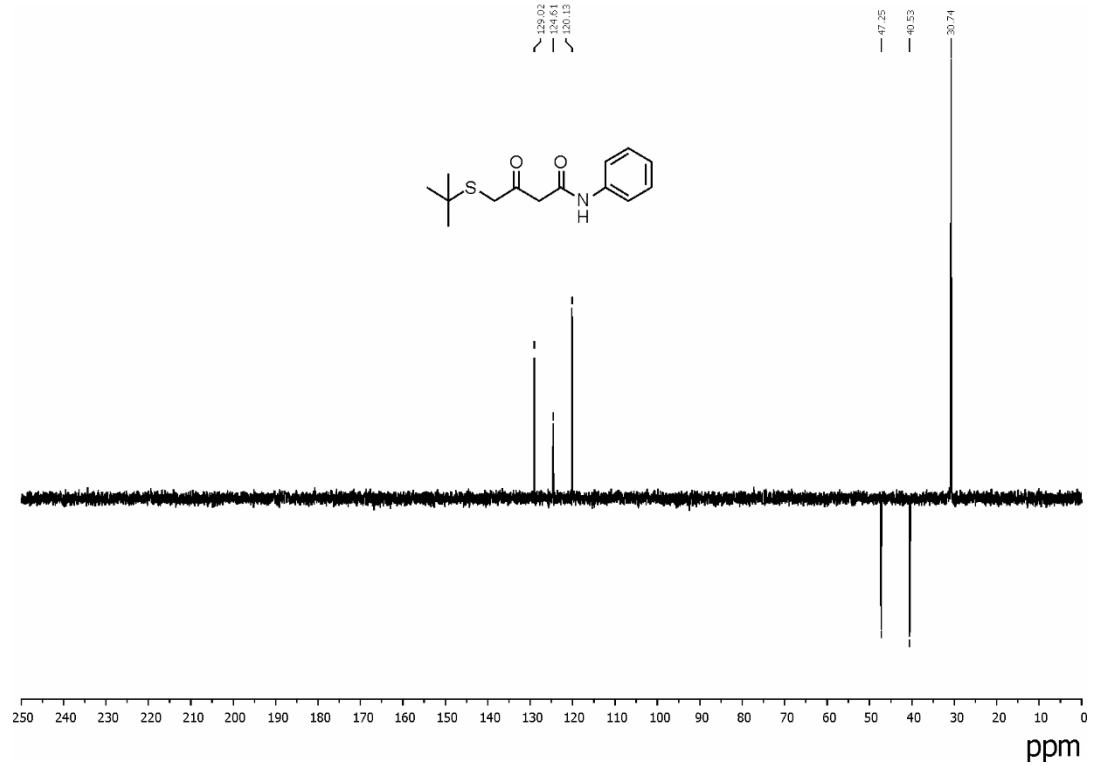

$^{13}$C NMR **3.027**



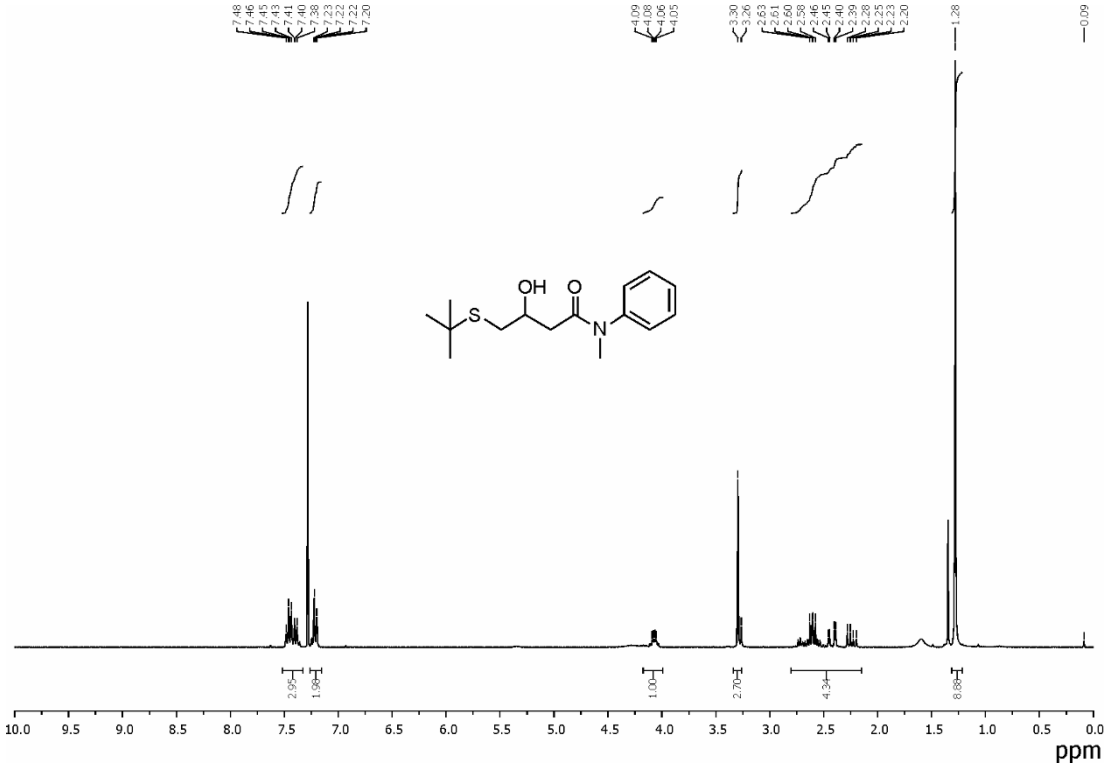
$^{1}$H NMR **3.028**



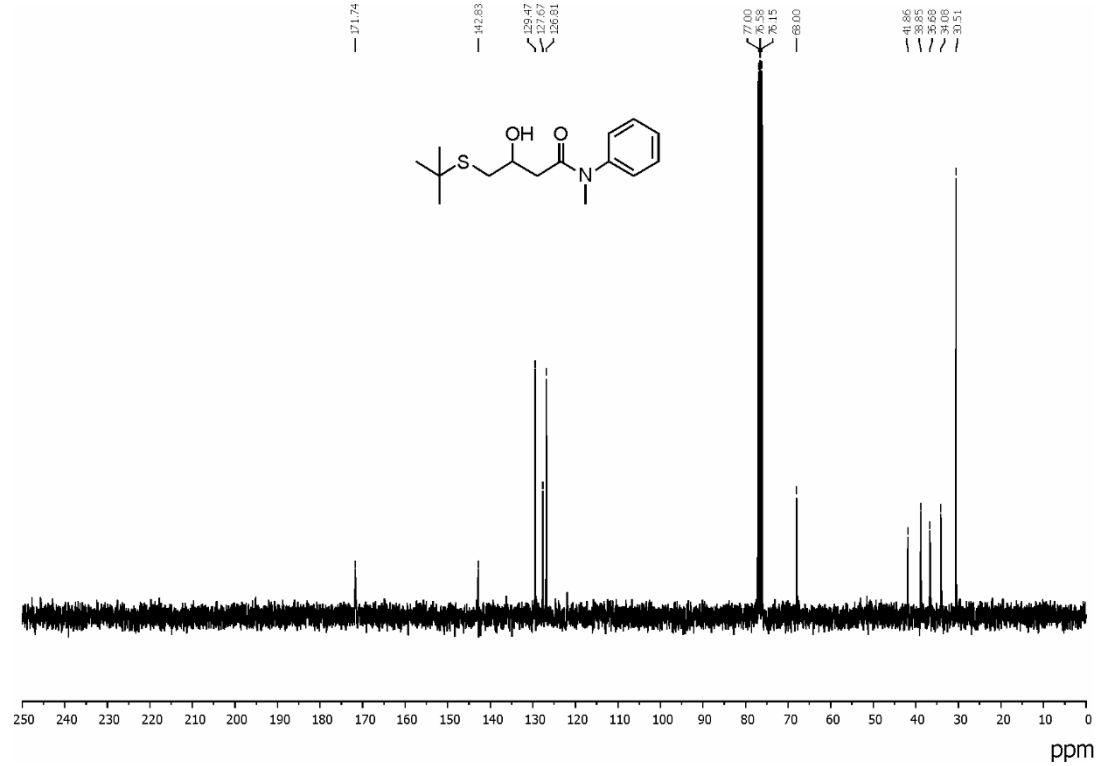

APT NMR **3.028**



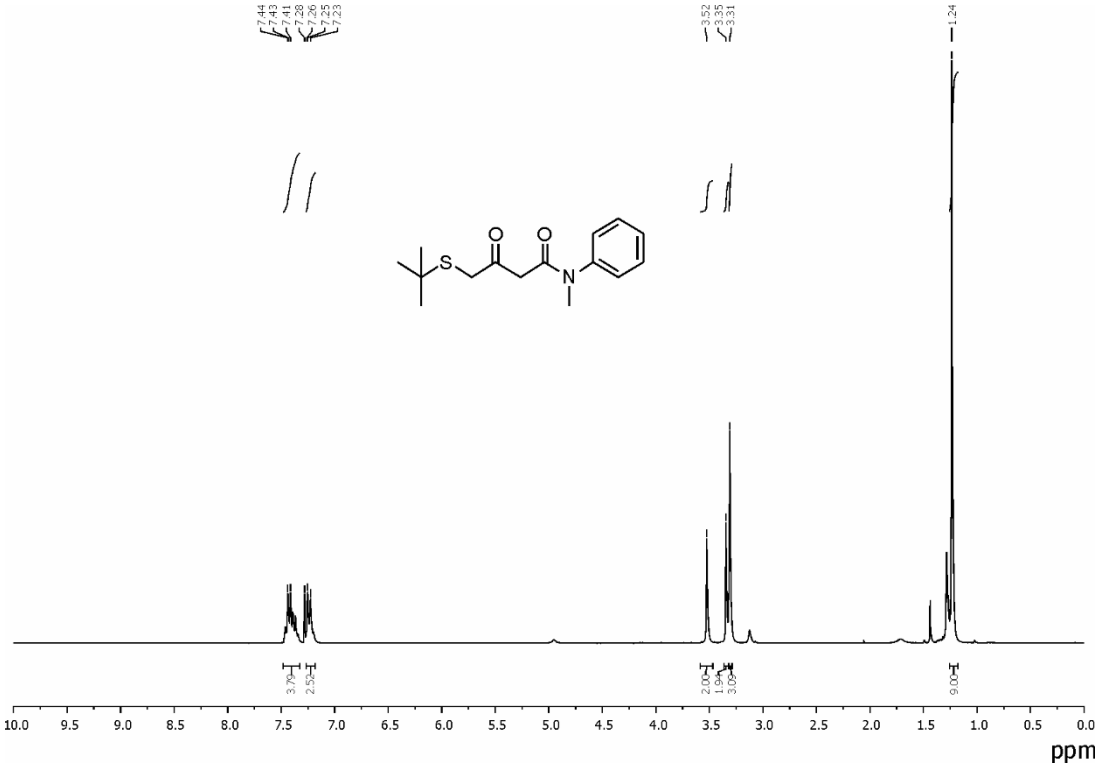
$^{1}$H NMR **3.029**



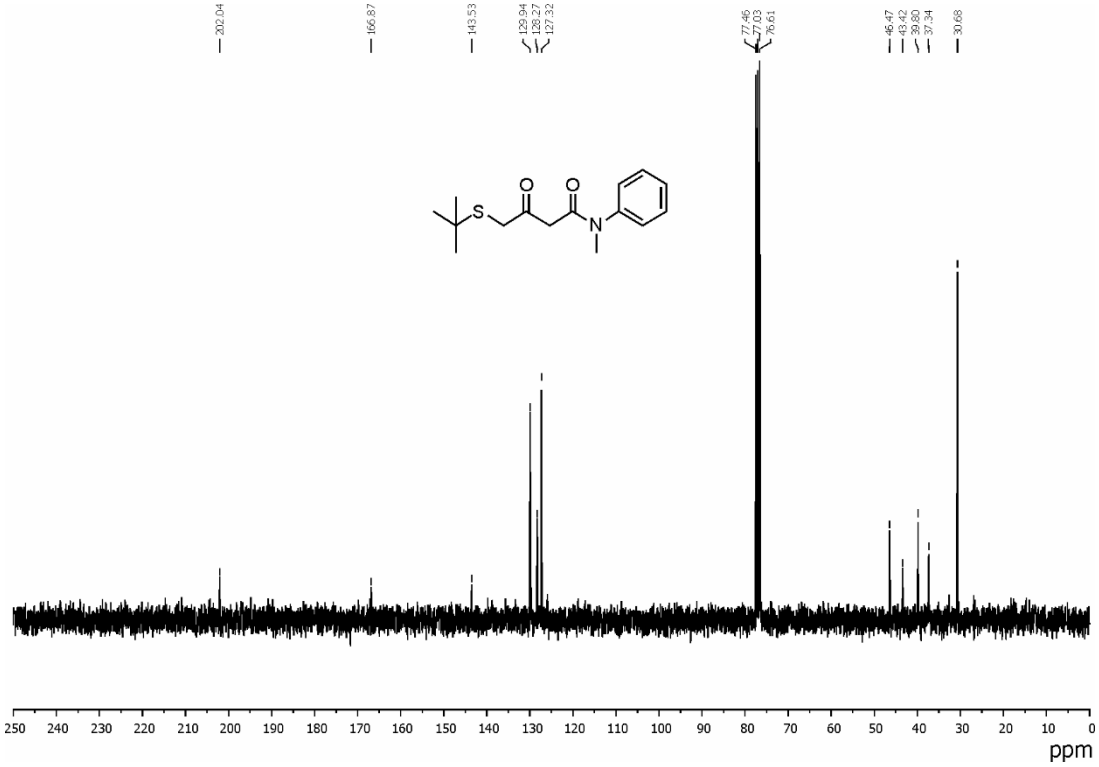

$^{13}$C NMR **3.029**



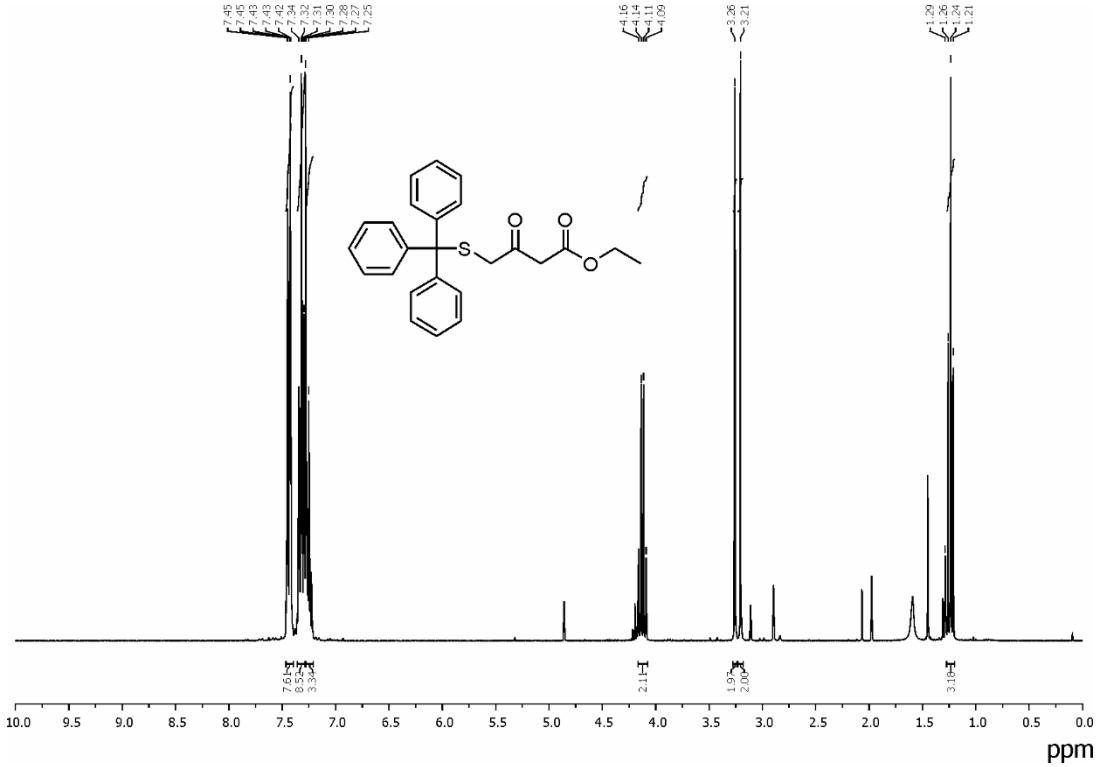
$^{1}$H NMR **3.031**



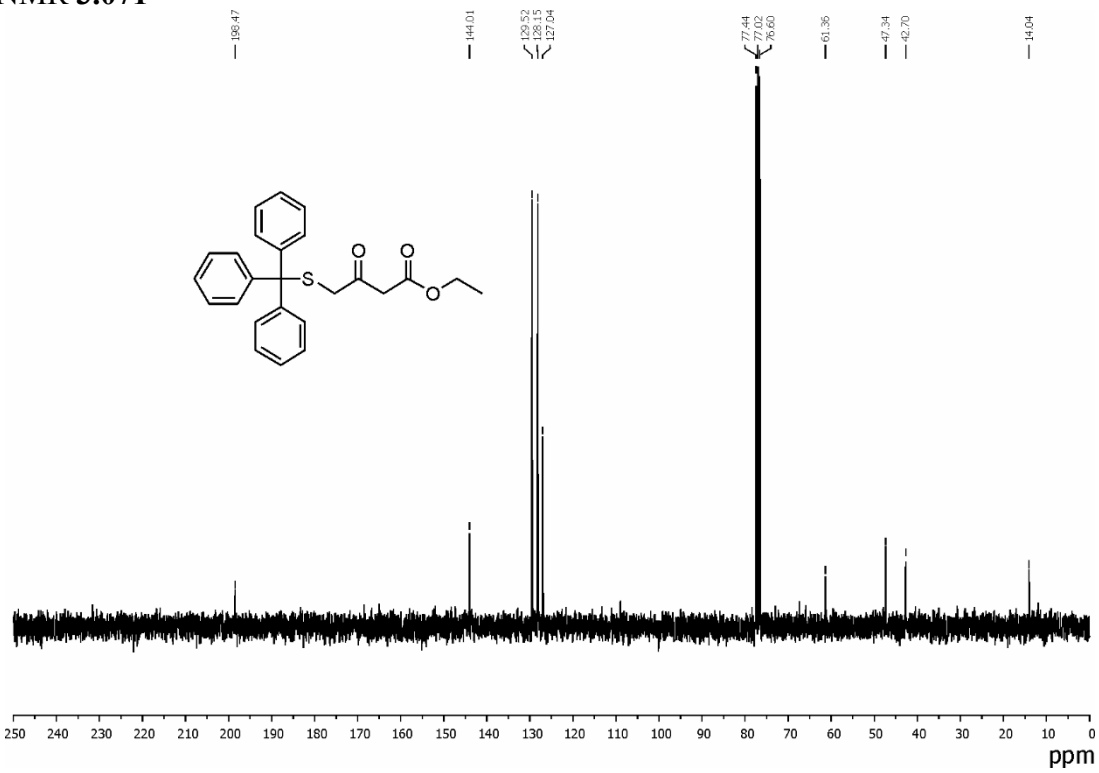

$^{13}$C NMR **3.031**



$^{1}$H NMR **3.071**

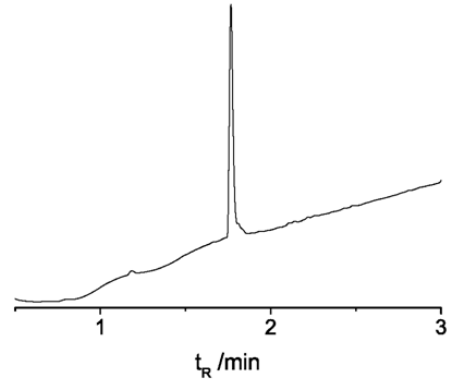


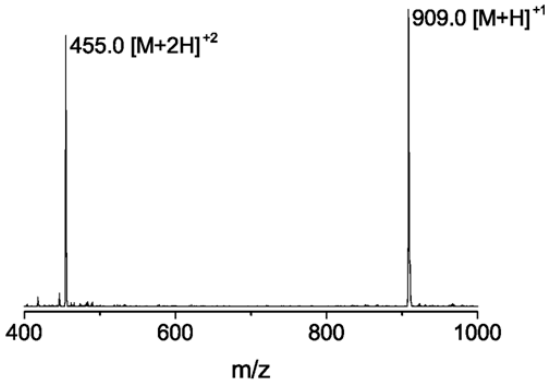

$^{13}$C NMR **3.071**



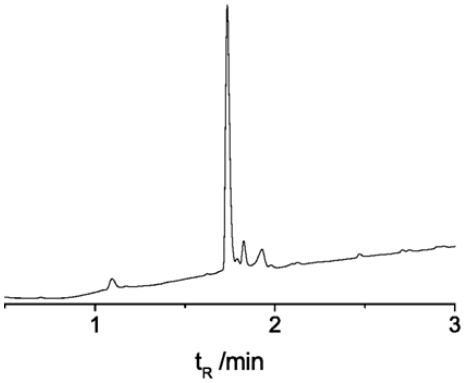
## A.2 UPLC-MS Analysis

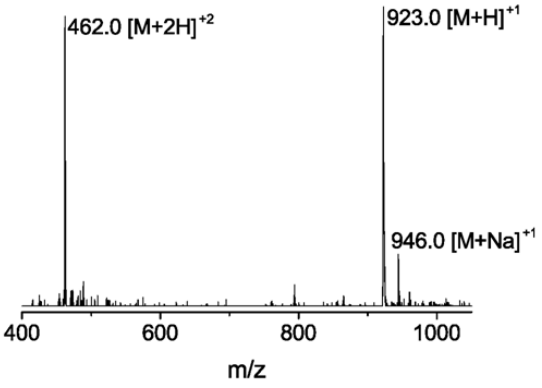
H-GRAEYSGLG-$NH_2$ (**3.032**)



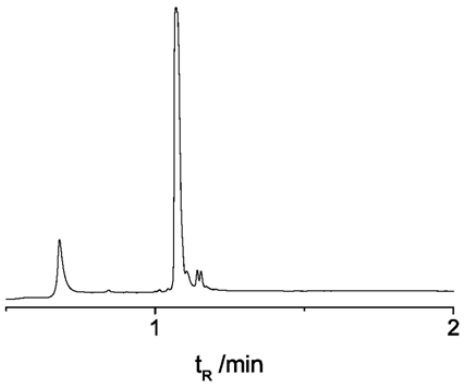



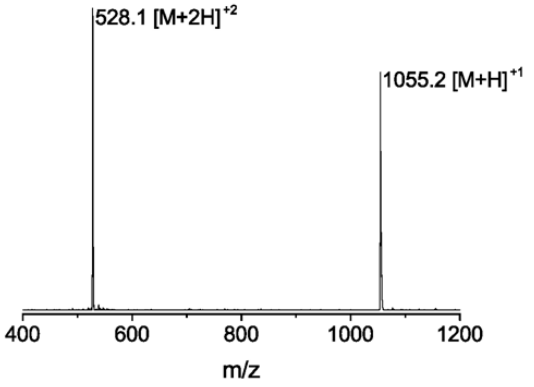

H-ARAEYSGLG-$NH_2$ (**3.033**)



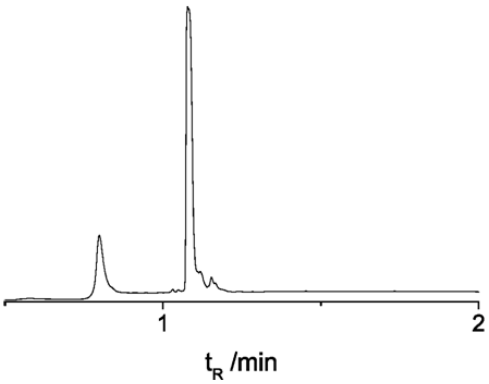



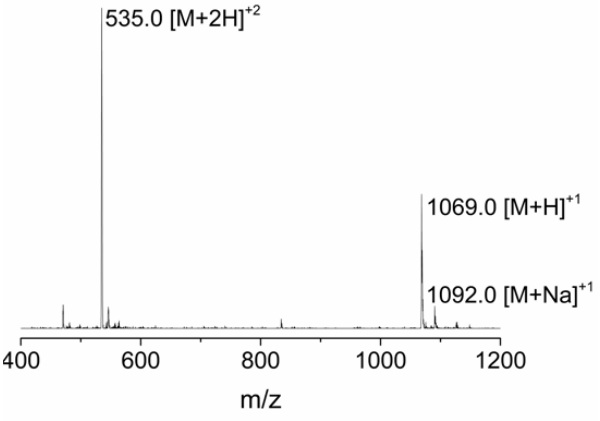

Aux-GRAEYSGLG-$NH_2$ (OEt) (**3.042**)



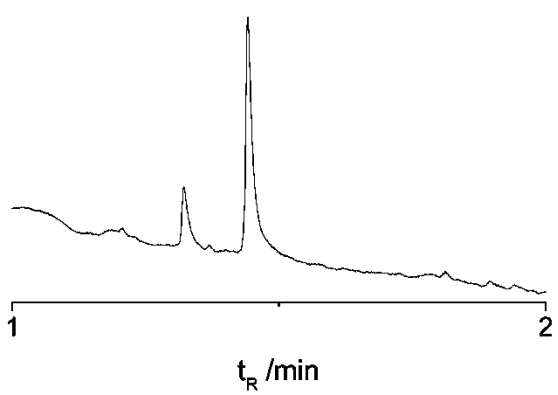



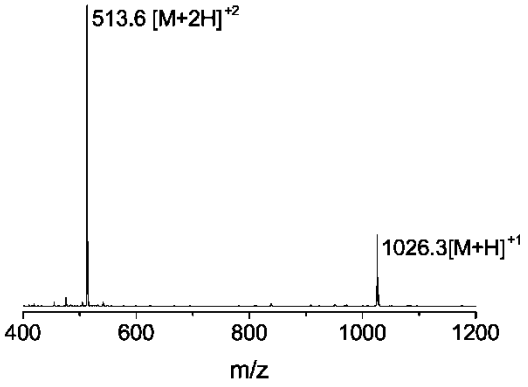

Aux-ARAEYSGLG-$NH_2$ (OEt) (**3.043**)



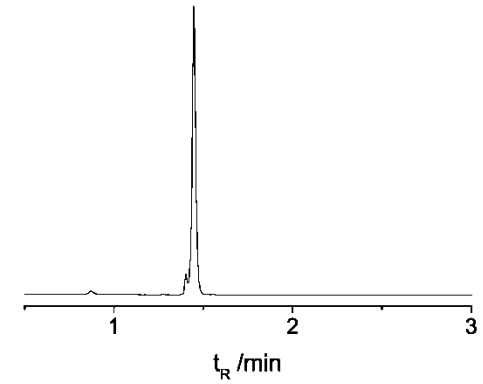



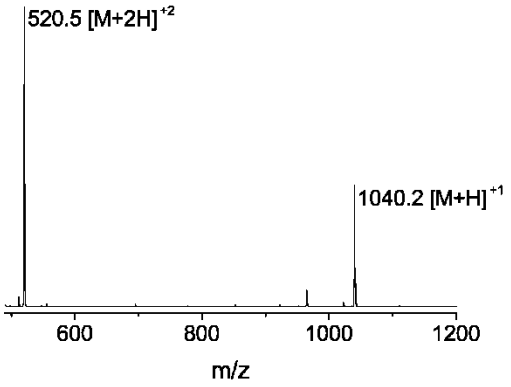
Aux-GRAEYSGLG-$NH_2$ ($NH_2$) (**3.044**)



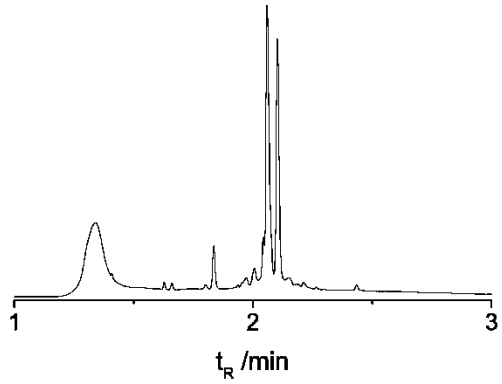



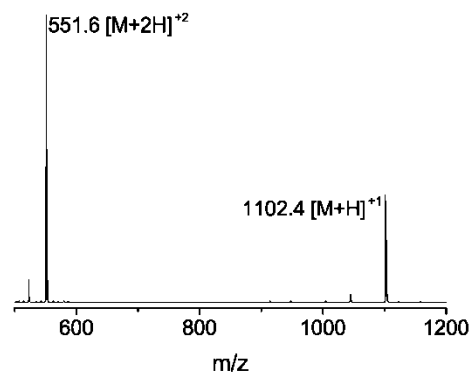

Aux-ARAEYSGLG-$NH_2$ ($NH_2$) (**3.045**)



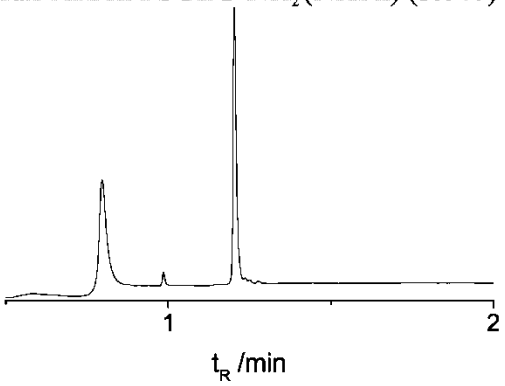



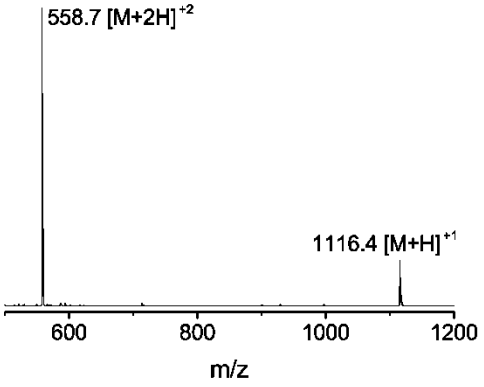

Aux-GRAEYSGLG-$NH_2$ (NHPh) (**3.046**)



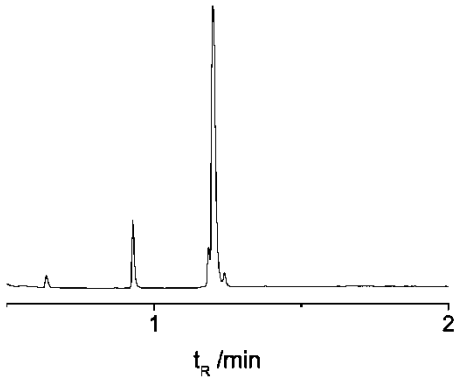



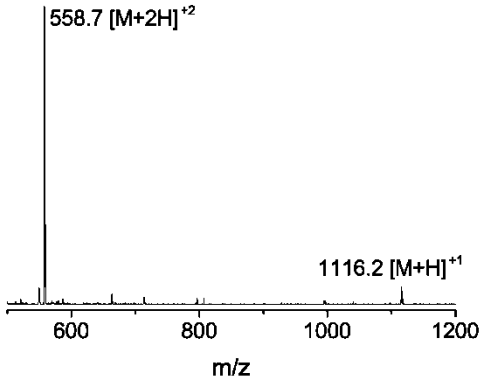

Aux-ARAEYSGLG-$NH_2$ (NHPh) (**3.047**)



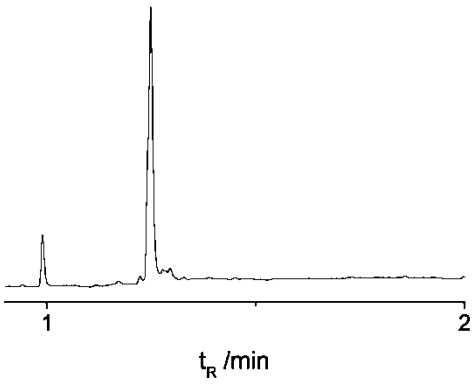



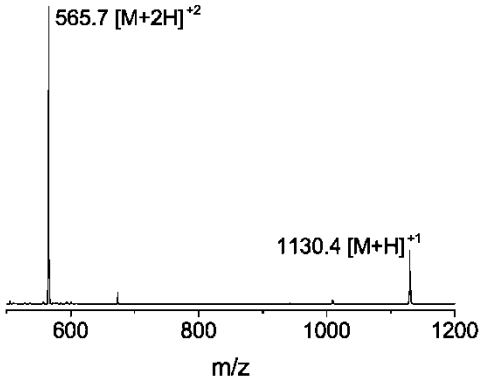
Aux-GRAEYSGLG-$NH_2$ (NMePh) (**3.048**)



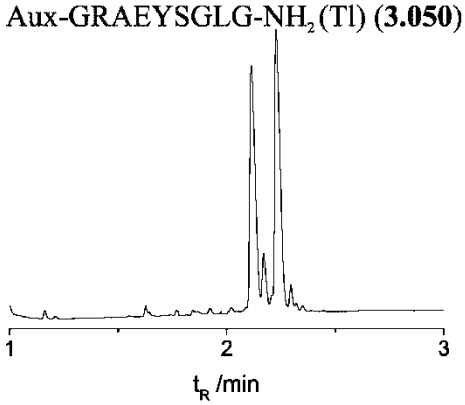



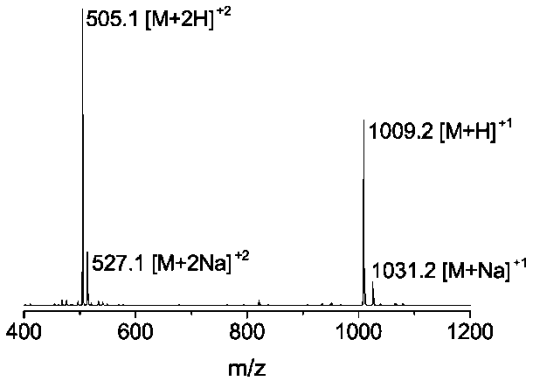

Aux-ARAEYSGLG-$NH_2$ (NMePh) (**3.049**)



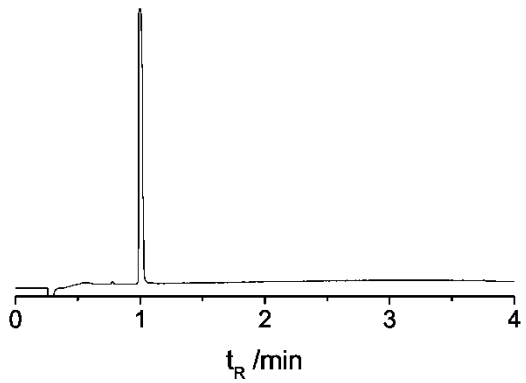



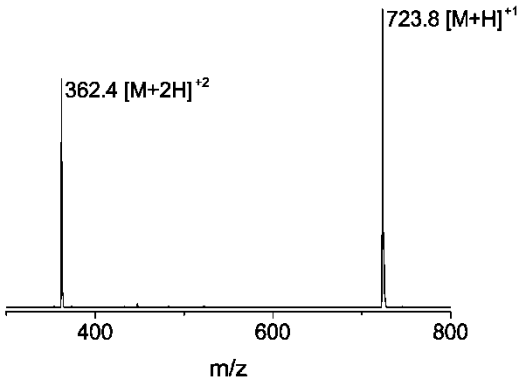

Aux-GRAEYSGLG-$NH_2$ (Tl) (**3.050**)



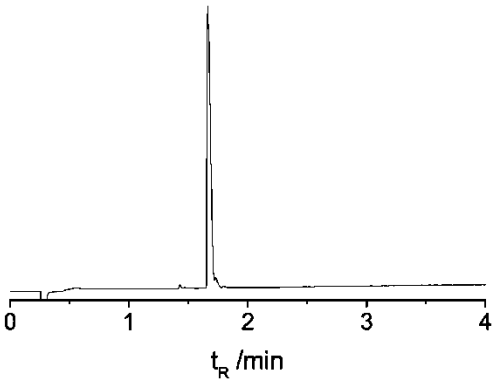



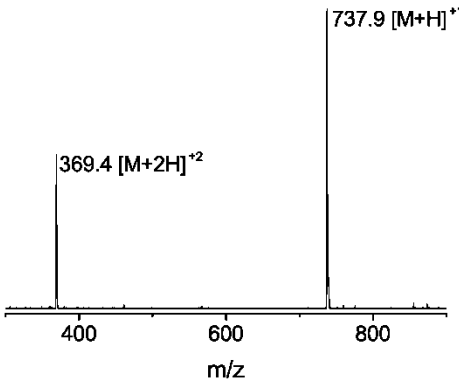

H-LYRAG-MPA-G-$NH_2$ (**3.052**)



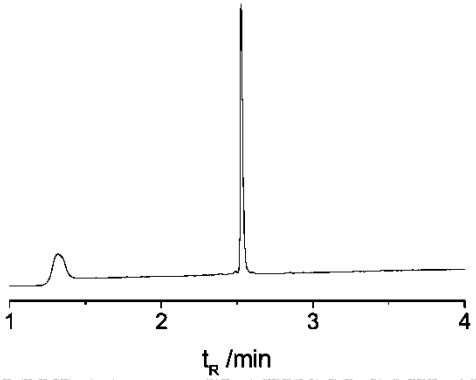



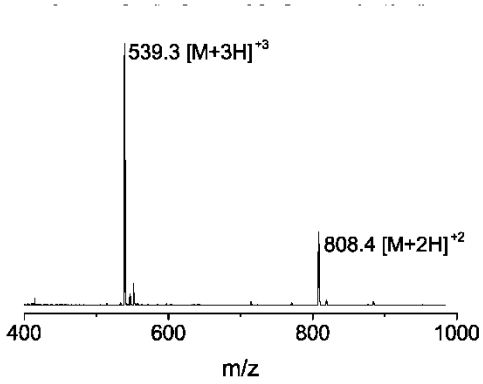
H-LYRAA-MPA-G-$NH_2$ (**3.053**)



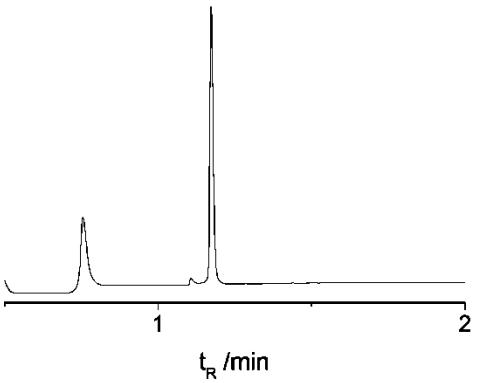



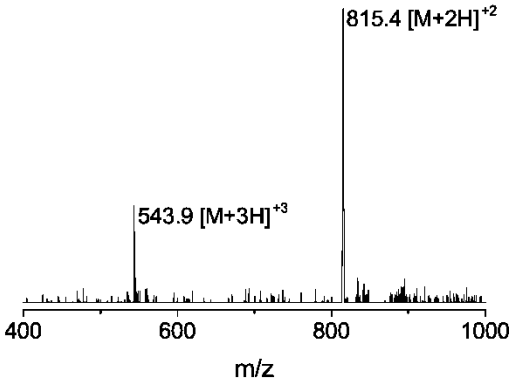

H-LYRAG-Aux-GRAEYSGLG-$NH_2$ (OEt) (**3.054**)



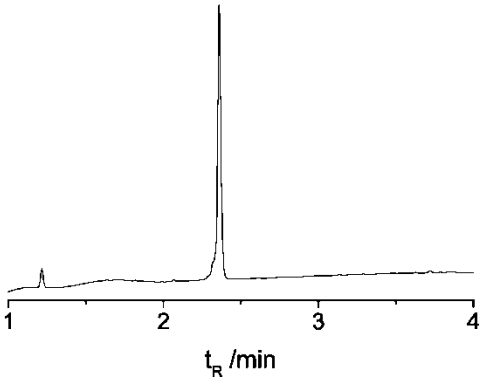



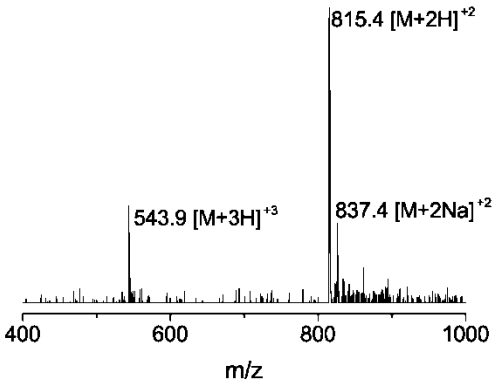

H-LYRAA-Aux-GRAEYSGLG-$NH_2$ (OEt) (**3.055**)



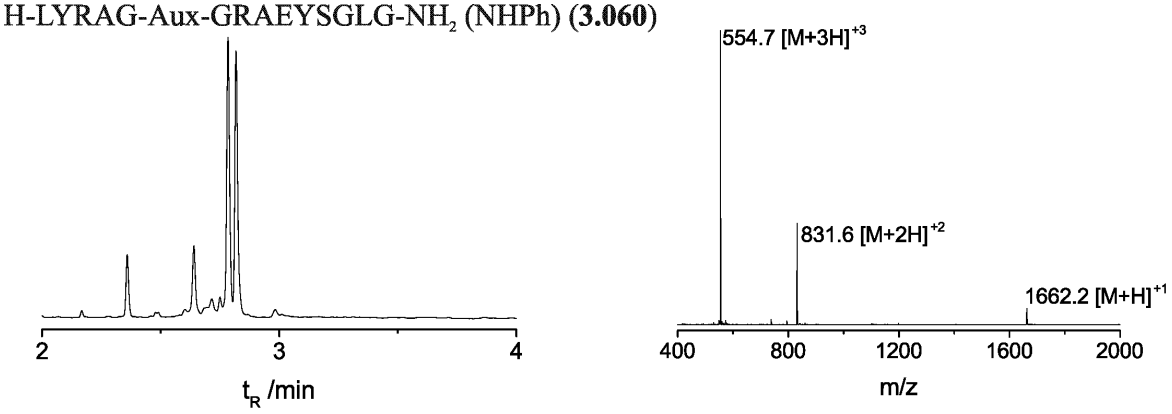




H-LYRAG-Aux-ARAEYSGLG-$NH_2$ (OEt) (**3.056**)

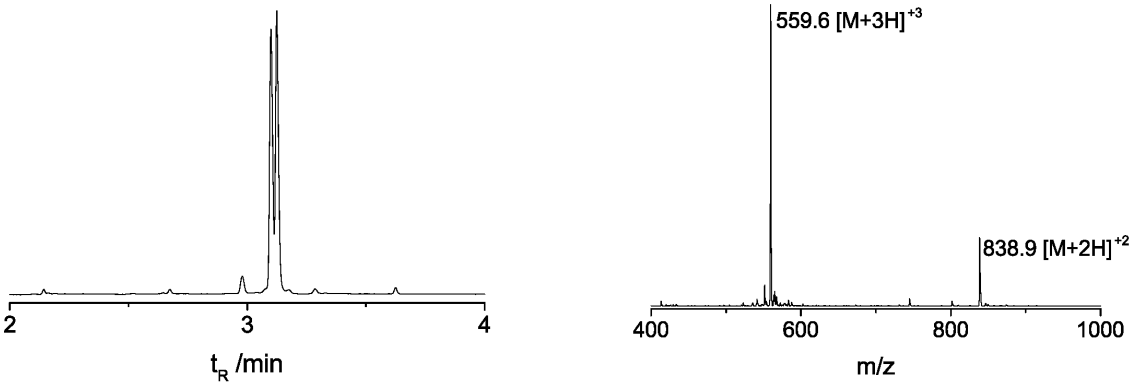


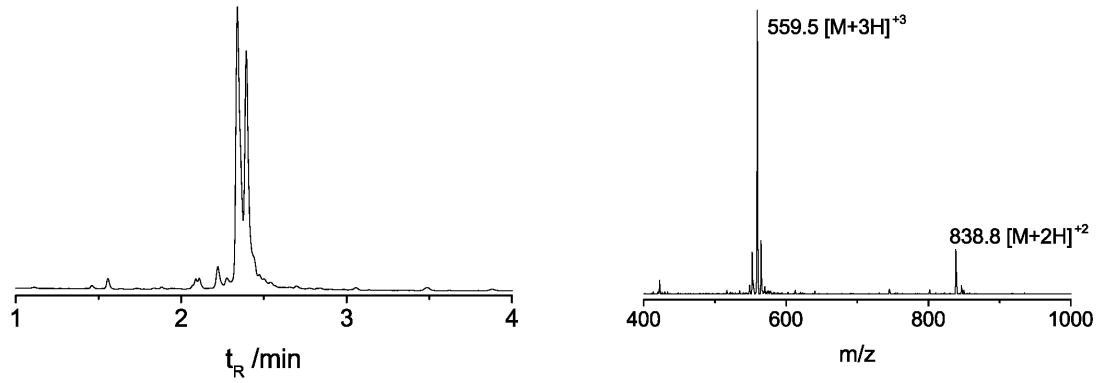
H-LYRAG-Aux-GRAEYSGLG-$NH_2$ (NHPh) **(3.060)**



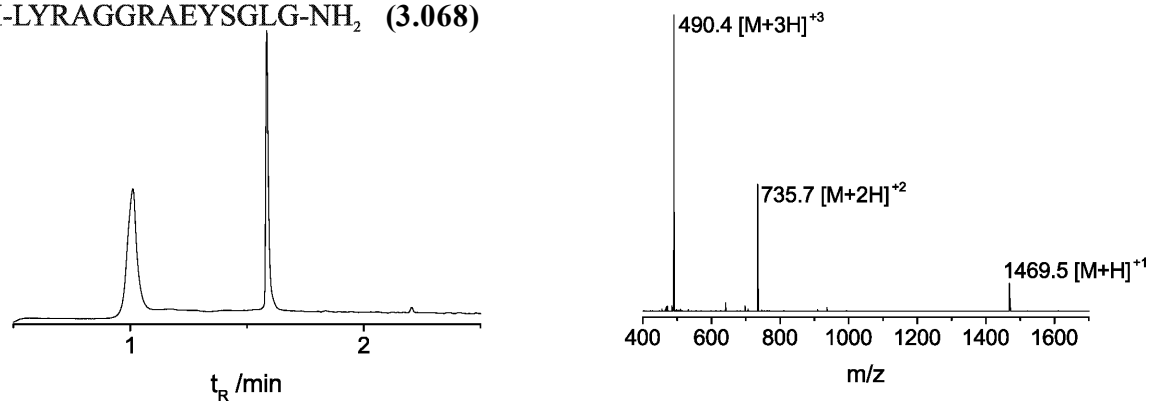

H-LYRAA-Aux-GRAEYSGLG-$NH_2$ (NHPh) **(3.061)**



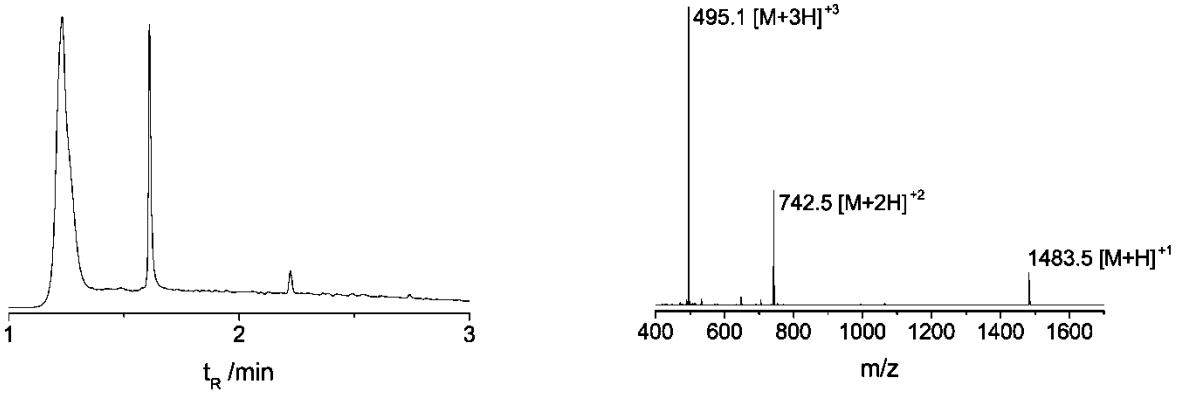

H-LYRAG-Aux-ARAYESGLG-$NH_2$ (NHPh) **(3.062)**



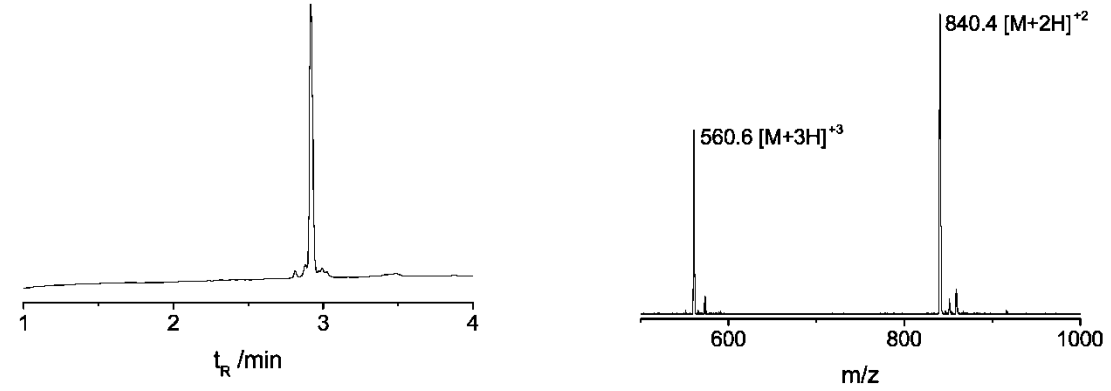

H-LYRAGGRAEYSGLG-$NH_2$ **(3.068)**



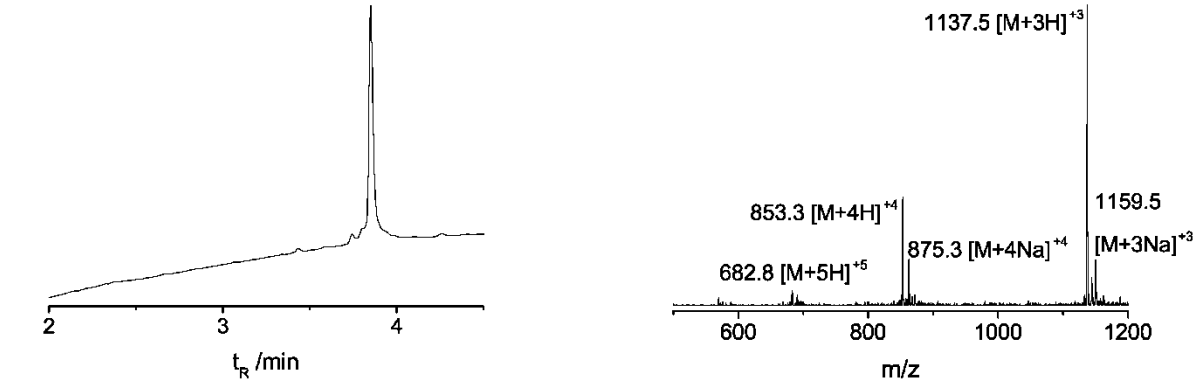
H-LYRAAGRAEYSGLG-$NH_2$ (OEt) (**3.069**)



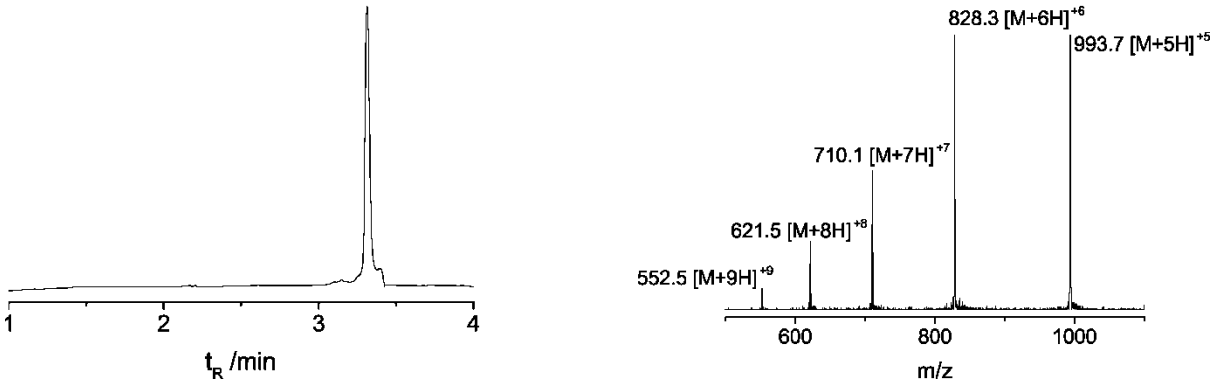

H-SSLLEKGLDGAKKAVG-SBzl (**Der1**)



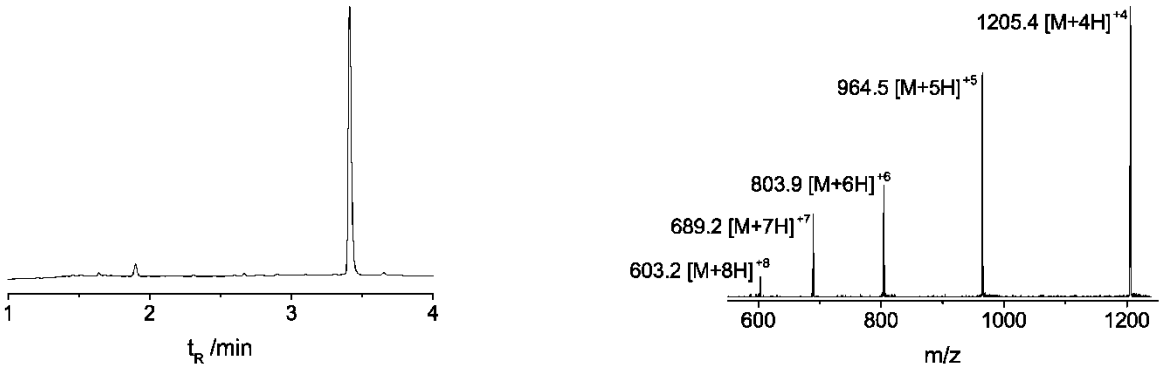

Aux-GLGKLGKDAVEDLESVGKGAVHDVKDVLDSVL-$NH_2$ (OEt) **(Der2)**



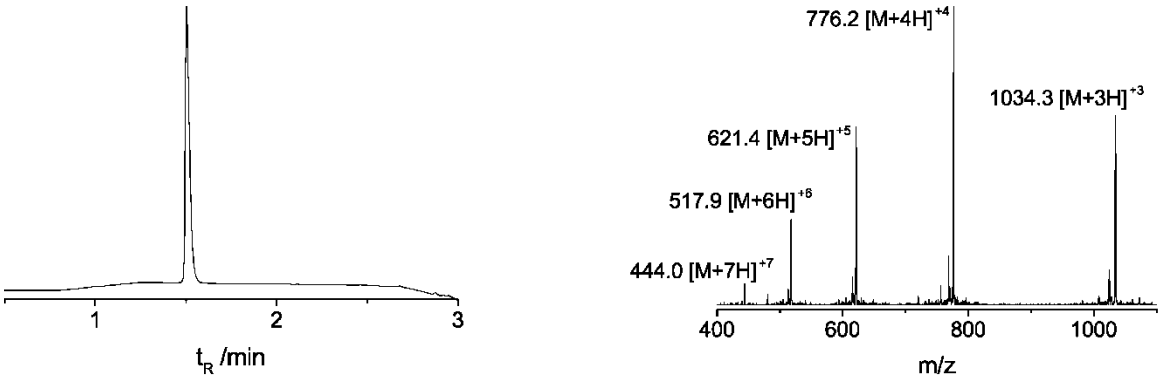

H-SSLLEKGLDGAKKAVG-Aux-GLGKLGKDAVEDLESVGKGAVHDVKDVLDSVL-$NH_2$ (OEt) **(Der12)**



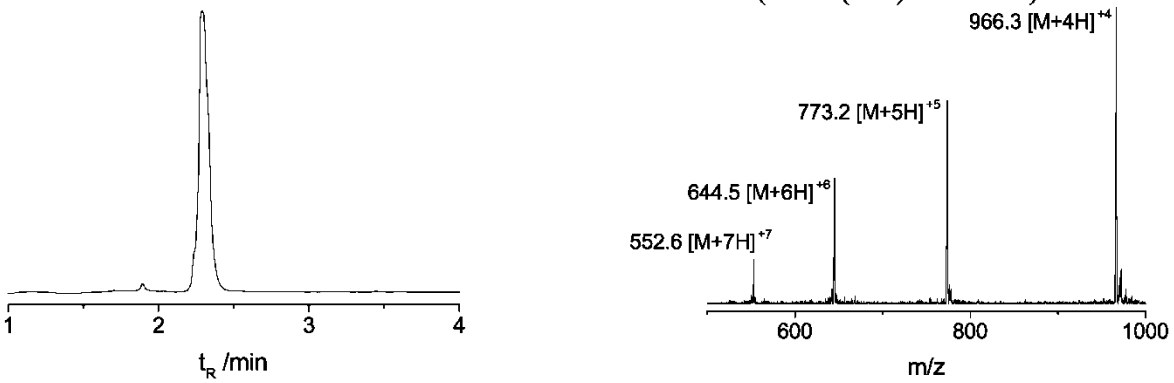
H-SSLLEKGLDGAKKAVGGLGKLGKDAVEDLESVGKGAVHDVKDVLDSVL-$NH_2$ **(DCD-1L)**



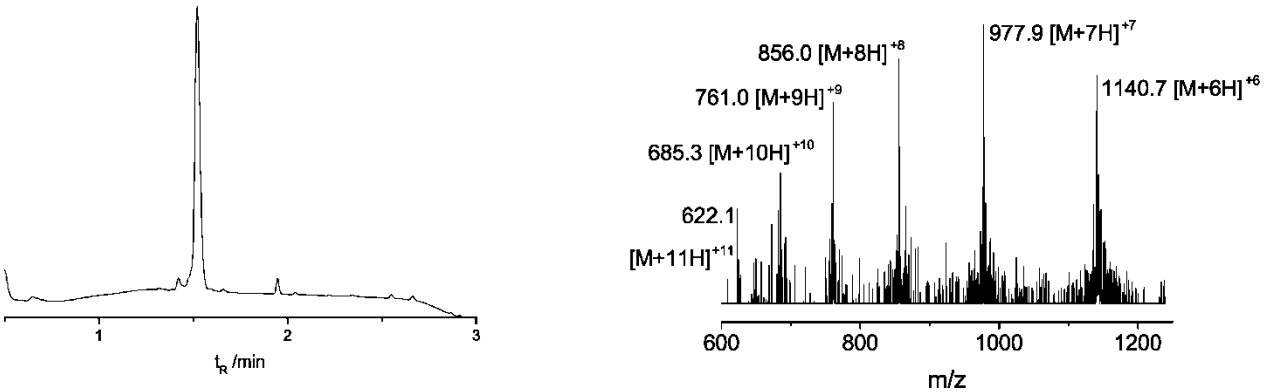

Aux-GVTSAPDTRPAPGSTAPPAHHHHHHH-$NH_2$ **(Aux-MUC1(1-20)-$His_6$ 3.079)**



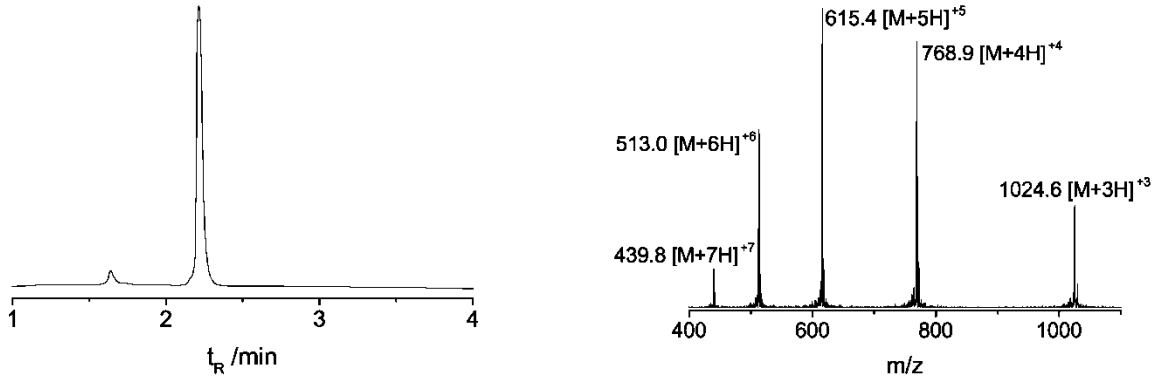

H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH-SBzl **(MUC1(1-40)-SBzl 3.080)**



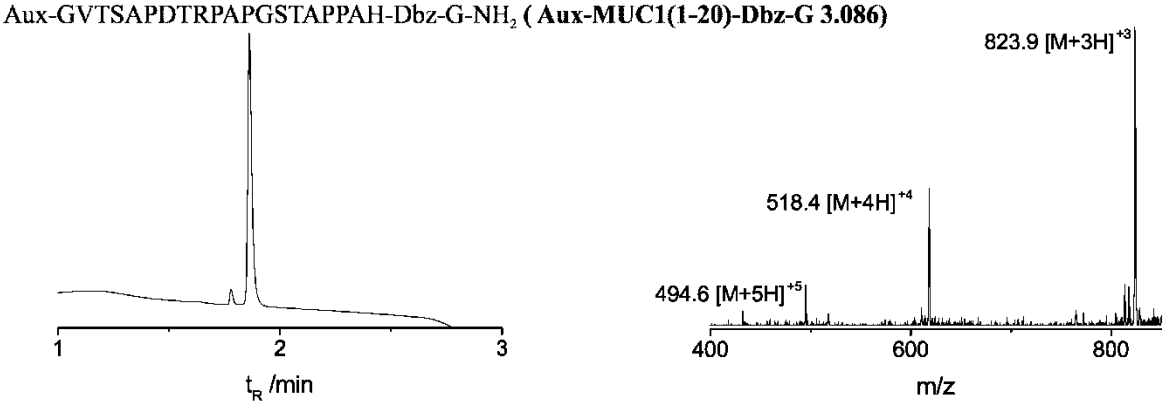

H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH-AUX- GVTSAPDTRPAPGSTAPPAH GVTSAPDTRPAPGSTAPPAHHHHHHH **(MUC1(1-40)-Aux-MUC1(1-20)-$His_6$ 3.081)**



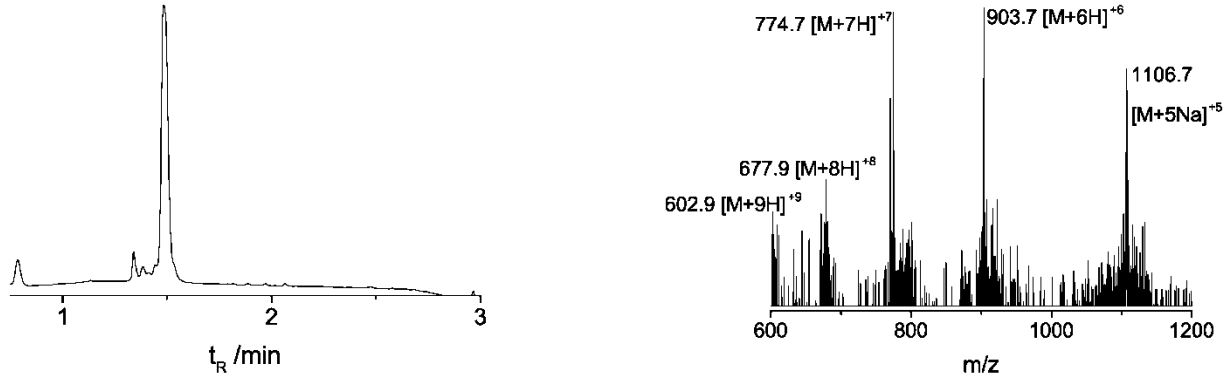
H-HHHHHHGVTSAPDTRPAPGSTAPPAH-SBzl **(His$_6$-MUC1(1-20)-SBzl 3.083)**



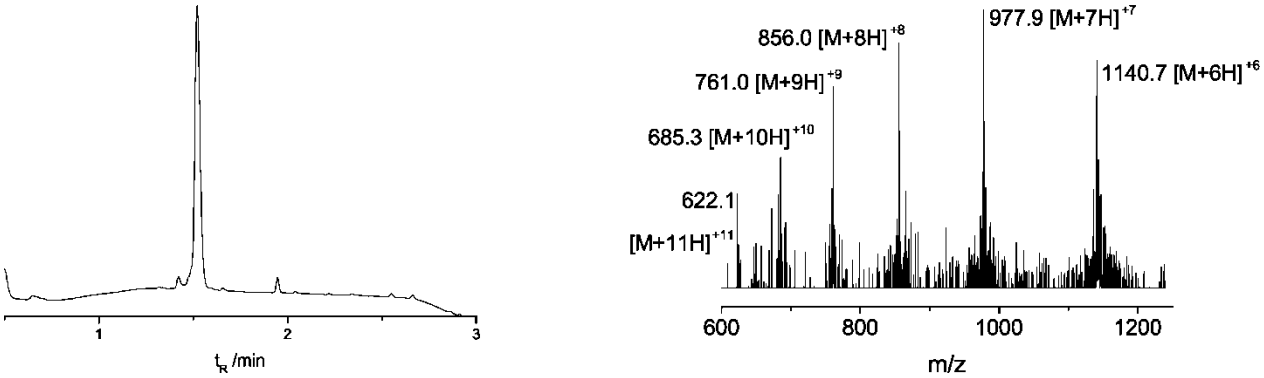

Aux-GVTSAPDTRPAPGSTAPPAH-Dbz-G-NH$_2$ **( Aux-MUC1(1-20)-Dbz-G 3.086)**



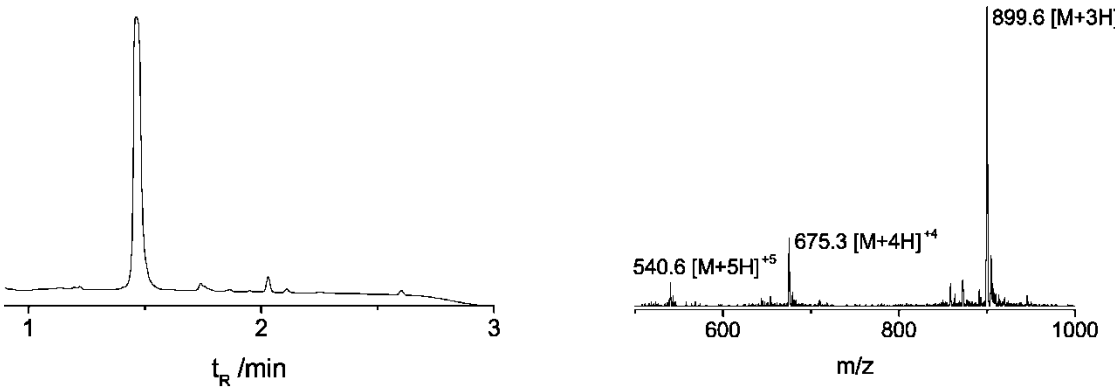

H-HHHHHHGVTSAPDTRPAPGSTAPPAH -Aux-GVTSAPDTRPAPGSTAPPAH-Dbz-G-NH$_2$
**(His$_6$-MUC1(1-20)-Aux-MUC1(1-20)-Dbz-G 3.095)**



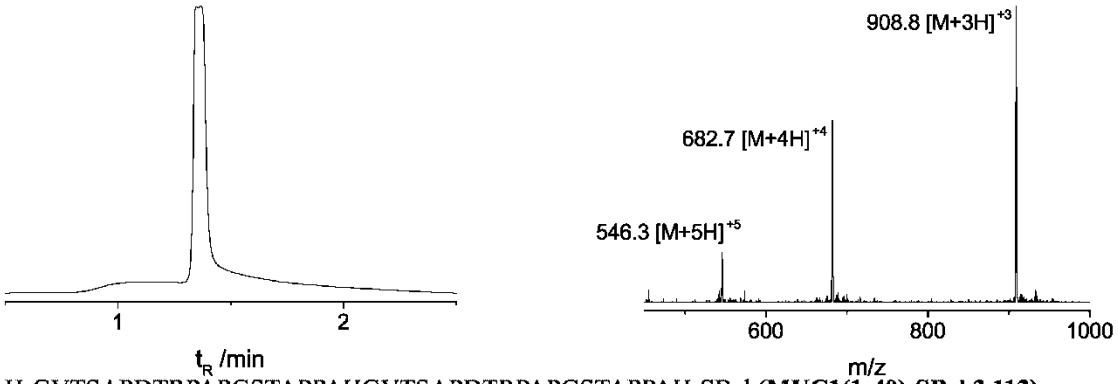

H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH-Aux-GVTSAPDTRPAPGSTAPPAH
HHHHHH-NH$_2$ **(MUC1(1-40)-Aux-MUC1(1-20)-His$_6$ 3.096)**



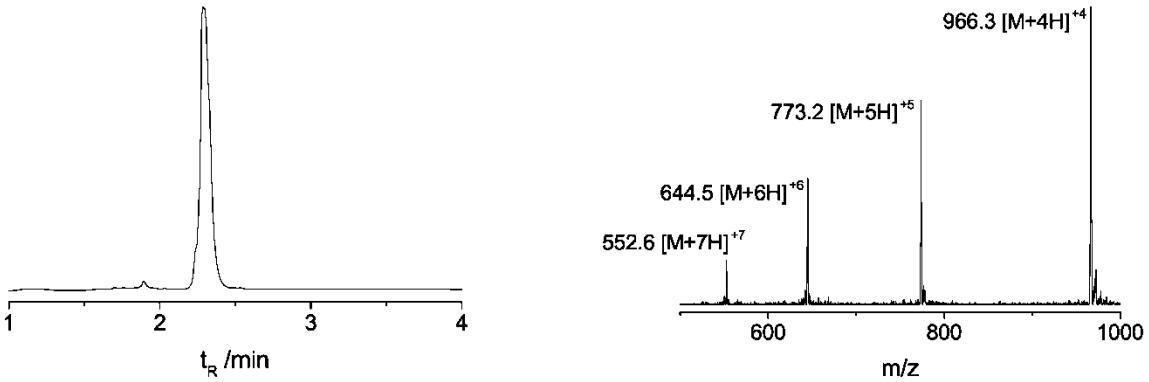
H-HHHHHGVTSAPDTRPAPGSTAPPAH-$SCH_2CH_2SO_3H$ **(His5-MUC1(1-20)-MESA 3.105c)**



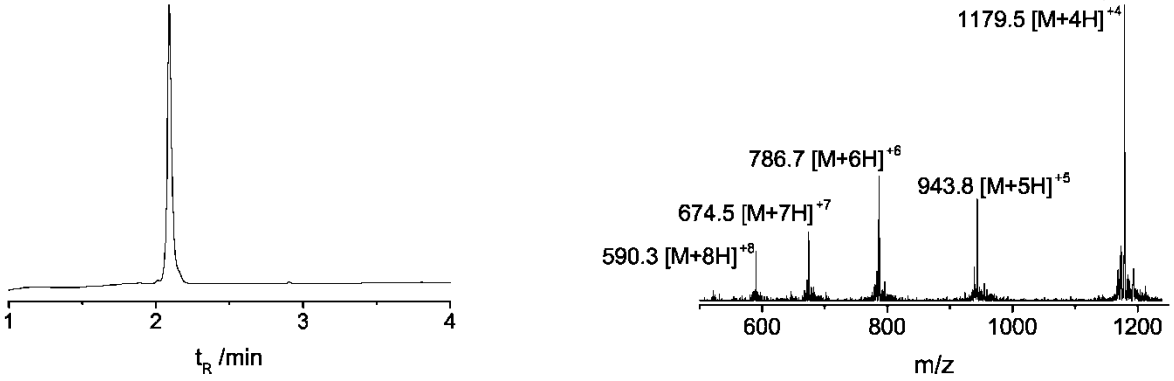

H-HHHHHHGVTSAPDTRPAPGSTAPPAH-$NHNH_2$ **(MUC1(1-20)-$NHNH_2$ (3.111)**



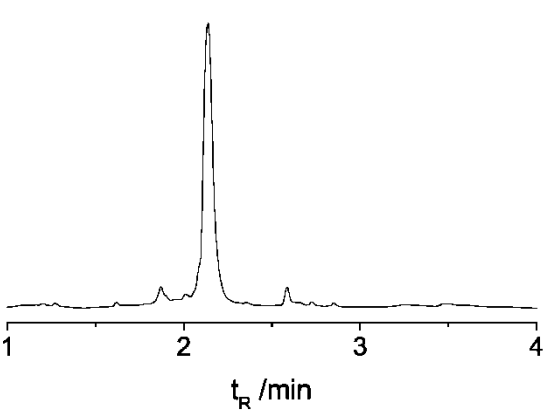

H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH-SBzl **(MUC1(1-40)-SBzl 3.112)**

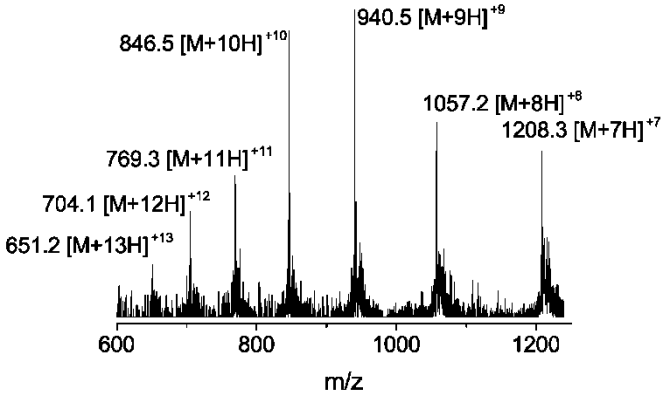



Aux-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHHHHHHH **(Aux-MUC1(41-80)-$His_6$ 3.114)**

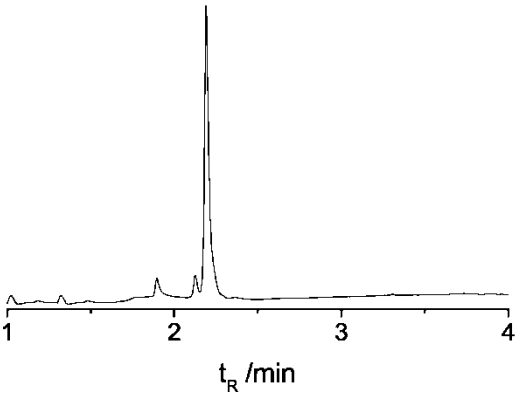


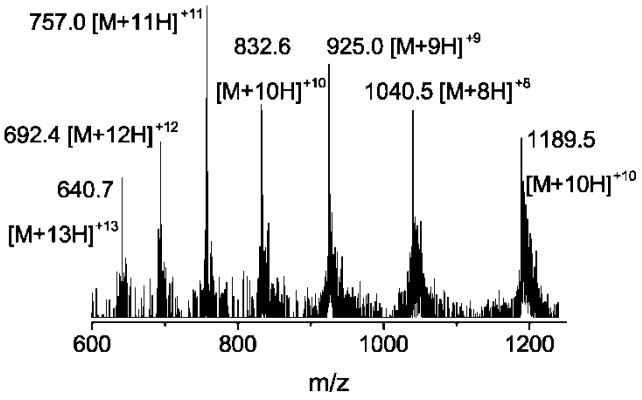
H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH-Aux-GVTSAPDTRPAPGSTAPPAH
GVTSAPDTRPAPGSTAPPAHHHHHHH-$NH_2$ **(MUC1(1-40)-Aux-MUC1(41-80)-$His_6$ 3.115)**




H-GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH
GVTSAPDTRPAPGSTAPPAHHHHHHH-$NH_2$ **(MUC1(1-80)-$His_6$ 3.116)**



# Appendix B: List of Abbreviations

| | |
|---|---|
| Ac | Acetyl |
| $Ac_2O$ | Acetic anhydride |
| Acm | Acetamidomethyl |
| Alloc | Allyloxycarbonyl |
| AMP | Antimicrobial peptide |
| Ar | Aryl |
| Aux | Auxiliary |
| Bnz | Benzyl |
| Boc | *tert*-Butoxycarbonyl |
| BOP | Bis(2-oxo-3-oxazolidinyl)phosphine |
| *t*-Bu or $^t$Bu | *tert*-Butyl |
| Bzl | Benzyl |
| cHex | Cyclohexane |
| CI2 | Chymotrypsin Inhibitor 2 |
| Da | Dalton |
| Dbz | 3,4-Dimethylbenzoic acid |
| DCD-1L | Dermicidin |
| DCM | Dichloromethane |
| DIC | Diisopropylcarbodiimide |
| DIPEA | Diisopropylethylamine |
| DMF | *N,N*-Dimethylformamide |
| DMP | Dess-Martin periodinane |
| DMSO | Dimethylsulphoxide |
| DPA | Diisopropylamine |
| EDC | 1-Ethyl-3-(3-dimethylaminopropy)carbodiimide |
| EDMA | Ethyldimethylamine |
| Eq | Equivalent |
| ESI | Electrospray ionization |
| Et | Ethyl |
| EtOH | Ethanol |
| $Et_2O$ | Diethyl ether |
| *et al.* | et alia |
| EtOAc | Ethyl Acetate |
| EWG | Electron withdrawing group |
| FMOC | 9-Fluorenylmethoxycarbonyl |
| Gn.HCl | Guanidinium hydrochloride |
| HATU | 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid |
| HBTU | 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate |
| HCTU | *O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium |
| HCY | Homocysteine |
| HPLC | High performance liquid chromatography |
| HOBt | 1-Hydroxybenzotriazole |
| hr | hour |
| KAHA | ketoacid-hydroxylamine |
| KCL | Kinetically controlled ligation |
| MBHA | 4-methylbenzhydrylamine |
| Me | Methyl |
| MeOH | Methanol |
| MESA | 2-Mercaptoethylsulfonic acid |
| MESNa | Sodium 2-mercaptoethylsulfonate |
| MPA | 3-mercaptopropylacetic acid |

| | |
|---|---|
| MPAA | 4-mercaptophenylacetic acid |
| MS | Mass spectroscopy |
| MUC1 | Mucin 1 |
| Nbz | 1H-benzo[d]imidazol-2(3H)-one |
| NCL | Native chemical ligation |
| NMA | N-methylaniline |
| NMM | N-Methylmorpholine |
| NMP | N-Methylpyrrolidone |
| NMR | Nucleic magnetic resonance |
| NTA | Nitriloacetic acid |
| OxymaPure | Ethyl (hydroxyimino)cyanoacetate |
| Ph | Phenyl |
| Pip | Piperidine |
| *i*-Pr or $^{i}$Pr | *iso*-propilo |
| PyBOP | benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate |
| rt | Room temperature |
| SAL | Sugar assisted ligation |
| BzlSH/ SBzl | benzylmercpatane |
| SCAL | Side chain assisted ligation |
| SEA | bis(2-sulfenyl)amino |
| SPCL | Solid phase chemical ligation |
| SPPS | Solid phase peptide synthesis |
| PhSH/ SPh | Thiophenol/yl |
| TCEP | *tris*(2-carboxyethyl)phosphine |
| TEA | Triethylamine |
| TES | Triethylsilane |
| TFA | Triflouroacetic acid |
| TFMSA | Trifluoromethanesulfonic acid |
| THF | Tetrahydrofurane |
| Thz | Thiazolidine |
| TIS | Triisopropylsilyl |
| TL | Thiolactone |
| Tmob | 4,5,6-trimethoxy-2-mercaptobenzyl |
| TMOF | trimethylorthoformate |
| TMS | Trimethylsilyl |
| $t_R$ | Retention time |
| Trt | Trityl, triphenylmethyl |
| UPLC | Ultra performance liquid chromatography |
| UV | Ultra violet |
| VNTR | variable number tandem repeat |